TAMPEREEN TEKNILLINEN YLIOPISTO
TAMPERE UNIVERSITY OF TECHNOLOGY

MIIA NIEMELÄ

COMPUTATIONAL STUDY OF G PROTEIN–COUPLED RECEPTOR PARTITIONING IN A PHASE SEGREGATED MEMBRANE

Master of Science Thesis

Examiner: Prof. Ilpo Vattulainen
Examiner and topic approved by the
Faculty of Natural Sciences Council
Meeting on September 7th 2016

ABSTRACT

MIIA NIEMELÄ: Computational Study of G Protein–Coupled Receptor Partitioning in a Phase Segregated Membrane
Tampere University of Technology
Master of Science Thesis, 77 pages, 12 Appendix pages
August 2016
Master's Degree Programme in Science and Engineering
Major: Technical Physics
Examiner: Prof. Ilpo Vattulainen
Supervisor: Dr. Moutusi Manna
Keywords: G protein-coupled receptor, protein partitioning, lipid phase separation, membrane domains, molecular dynamics

Cell membranes are composed of a great diversity of lipids and proteins. The lipid raft concept suggests that instead of a homogeneous distribution of the membrane components, they are laterally segregated into domains of compositional and functional diversities. The rafts are thought to be functional membrane domains rich in saturated lipids and cholesterol, serving as platforms for signaling and lateral sorting of membrane components.

G protein–coupled receptors (GPCRs) are versatile signaling machines that control a large variety of physiological processes. For example, the cyclic adenosine monophosphate (cAMP) signaling pathway used for cellular communication is triggered by GPCRs. Many signaling components in the cAMP pathway have been associated to rafts. The compartmentalization of signaling components to small regions would enable a rapid and accurate activation of the signaling pathways. It is not clear, however, whether GPCRs themselves reside in raft-like domains. GPCRs constitute the largest class of integral membrane proteins and are major drug targets. Therefore, further knowledge on their location in cell membranes could be significant in understanding the signaling processes they trigger and furthermore in the development of drugs acting on them.

The aim of this work is to investigate the partitioning of the β_2-adrenergic receptor, a characteristic GPCR, in a phase segregated membrane by molecular dynamics simulations, and thus to find out the receptor's preference for different domains. This is done by systematically placing the receptor in different membrane domains, such as liquid-ordered (L_o) or liquid-disordered (L_d) domains, or in a random distribution of a phase separating lipid mixture. According to the results, the receptor does not completely reside either in the raft-like L_o domain or in the L_d domain. The receptor repeatedly seeks to the L_o/L_d boundary irrespective of the initial location, and creates a special environment around it consisting of cholesterol and L_d-associated unsaturated lipids.

TIIVISTELMÄ

MIIA NIEMELÄ: Laskennallinen tutkimus G-proteiinikytkentäisen reseptorin partitiosta faasijakautuneessa solukalvossa
Tampereen teknillinen yliopisto
Diplomityö, 77 sivua, 12 liitesivua
Elokuu 2016
Teknis-luonnontieteellinen koulutusohjelma
Pääaine: Teknillinen fysiikka
Tarkastaja: Prof. Ilpo Vattulainen
Ohjaaja: FT Moutusi Manna
Avainsanat: G-proteiinikytkentäinen reseptori, proteiinin partitio, lipidien faasijakautuminen, solukalvodomainit, molekyylidynamiikka

Solukalvo koostuu laajasta kirjosta erilaisia lipidejä ja proteiineja. Lipidilauttateoria ehdottaa, että näiden komponenttien satunnaisen jakautumisen sijaan ne jakautuisivatkin solukalvossa lateraalisesti koostumukseltaan ja toiminnaltaan erilaisiin faaseihin. Lipidilauttojen uskotaan olevan runsaasti tyydyttyneitä lipidejä ja kolesterolia sisältäviä toiminnallisia alueita, jotka muodostavat alustan solukalvokomponenttien viestinvälitykselle.

G-proteiinikytkentäiset reseptorit osallistuvat moniin fysiologisiin prosesseihin ja ovat tärkeitä solun viestinvälityksessä. Ne aktivoivat adenosiinimonofosfaattiväylän, jonka monien viestinvälityskomponenttien uskotaan sijaitsevat lipidilautoilla. Komponenttien keskittyminen juuri lipidilauttojen kaltaisille pienille alueille mahdollistaisi viestinvälityspolkujen nopean ja tarkan aktivoitumisen. On kuitenkin epäselvää, sijaitsevatko itse G-proteiinikytkentäiset reseptorit lipidilauttojen kaltaisilla alueilla. Suuri osa olemassa olevista lääkkeistä toimii G-proteiinikytkentäisten reseptorien kautta, ja siksi lisätieto niiden sijainnista solukalvoissa olisi arvokasta niiden aktivoimien prosessien ymmärtämisessä ja edelleen niiden kautta toimivien lääkkeiden kehityksessä.

Tämän työn tarkoituksena on tutkia molekyylidynamiikkasimulaatioiden avulla β_2-adrenergisen reseptorin, tyypillisen G-proteiinikytkentäisen reseptorin partitioitumista faasijakautuneessa kalvossa. Ongelmaa lähestytään sijoittamalla reseptori systemaattisesti erilaisiin faaseihin, kuten järjestyneisiin ja epäjärjestyneisiin alueisiin, sekä satunnaisesta lipidien jakaumasta muodostuviin lipidikalvoihin. Saatujen tulosten perusteella reseptori ei sijaitse täysin kummassakaan faasissa, vaan sen sijaan hakeutuu toistuvasti näiden kahden faasin rajalle alkusijainnistaan riippumatta, muodostaen ympärilleen erityisen kolesterolista ja tyydyttymättömistä lipideistä muodostuvan ympäristön.

PREFACE

This Master of Science Thesis was carried out in the Biological Physics and Soft Matter group of the Department of Physics at Tampere University of Technology between January 2016 and August 2016. The computing services for this study were provided by the Tampere Center for Scientific Computing (TCSC) and the Finnish IT Center for Scientific Computing (CSC).

I would like to thank my examiner Prof. Ilpo Vattulainen for the opportunity to work in this excellent group within intriguing projects, including the one behind this Thesis. I can honestly say that working in this group alongside my studies has taught me at least as much as the courses I have attended during these years. A special thanks goes to my supervisor Moutusi Manna for advising and supporting me throughout this project, even from long distance. I appreciate her expertise and warm guidance. I would also like to thank other members of our group for help and interesting discussions, especially the people in Batcave: Sami, Emmi, and Heikki, the days get never boring with you.

Finally, I would like to thank my family and friends for the encouragement and support during my studies and the writing process of this Thesis. The moments spent with the most important people have served as excellent breaks and given new boost to the writing process.

Tampere, August 2016

Miia Niemelä

TABLE OF CONTENTS

ABBREVIATIONS

AA	All-atom
AR	Adrenergic receptor
ATP	Adenosine triphosphate
β_2AR	β_2-adrenergic receptor
cAMP	Cyclic adenosine monophosphate
CHOL	Cholesterol
CG	Coarse–grained
COM	Center of mass
DLiPC	Dilinoleoyl–phosphatidylcholine
DOPC	Dioleoyl–phosphatidylcholine
DRM	Detergent resistant membrane
DPPC	Palmitoyl–phosphatidylcholine
GDP	Guanosine diphosphate
GPCR	G protein–coupled receptor
G–protein	Guanine nucleotide-binding protein
GTP	Guanosine triphosphate
GUV	Giant unilamellar vesicle
L_d	Liquid–disordered
L_o	Liquid–ordered
MD	Molecular dynamics
NMR	Nuclear magnetic resonance
NpT	Isobaric–isothermal ensemble
OPLS	Optimized potentials for liquid simulations
PC	Phosphatidylcholine
PBC	Periodic boundary conditions
PME	Particle–mesh Ewald
VMD	Visual molecular dynamics

1. INTRODUCTION

The world is made up of atoms: the plants, the rivers, the buildings and clothes, as well as our bodies are built from them. Many vitally important and scientifically interesting processes occur on the level of atoms, at the length scale of nanometers. Experimental research of nanoscale processes is challenging due to their small size and short duration, yet there are multiple research fields that are focused on atomic scale phenomena. Designing new materials and development of drugs, for example, require accurate knowledge of the physical properties that are too fast for human eye or beyond the limits of the current measurement devices. One solution to study systems on nanoscale is to use molecular simulations by describing complex chemical systems in terms of atomic models.

The molecular dynamics (MD) simulation technique is based on classical mechanics and provides a method for modeling nanoscale systems. It is a popular and powerful tool in the study of biological processes, such as membrane receptor mediated signal transduction. Several diseases are caused by disorders in membrane receptor function and can be controlled by drugs acting on the receptors, which has boosted the research of receptors and their function. The field is constantly developing; the computational power is growing, the methodologies are improving, and new high-resolution structures of membrane proteins are being resolved by X-ray crystallography. This has allowed the simulation of many structurally known membrane proteins in their native environment for durations of several microseconds. Due to the ability of MD simulation to reach atomic level description not accessible for most of the experimental techniques, the understanding of membrane protein function has taken major steps forward. MD simulations can reveal the dynamic behavior hidden in the static structures. Currently, a typical state-of-the-art simulation system includes several membrane proteins embedded to a patch of a lipid bilayer, allowing the investigation of lipid-lipid and protein-lipid interaction processes.

Cell membranes are formed of a large diversity of lipids and proteins that are designed to perform functions essential for the cell. Instead of being structurally homologous, the membrane is believed to control these functions by laterally segregating its components into lateral domains. In the study of model membranes, cholesterol–

containing lipid mixtures of unsaturated and saturated lipids are known to display phase coexistence over a broad range of compositions and temperatures [1, 2]. The lipids separate into a thicker, highly ordered, cholesterol–rich liquid-ordered (L_o) phase and into a thinner, more mobile, cholesterol–poor liquid-disordered (L_d) phase. The phase coexistence has been observed in studies of model membranes consisting of synthetic lipids [3–5], however it is debated whether they exist in live cell membranes since they are optically unresolvable. The lipid raft hypothesis proposes that these lipid domains are important for cellular function, such as signaling, trafficking, and lateral sorting of membrane components [6].

The G protein–coupled receptors (GPCRs) are important signal-mediators in cell membranes and form the largest family of human integral membrane proteins with over 800 members [7]. They are involved in a wide range of physiological processes related to several diseases, which has made them an attractive target for the development of drugs. Currently it is estimated that 30-50 % of the marketed drugs act on GPCRs [8], for which reason the studies of the GPCRs' structure, function, and regulation are at the forefront of the research today. The importance of the GPCR–related research was highlighted by the Nobel Prize awarded in 2012 for the discovery of the X–ray structure of the β_2–adrenergic receptor (β_2AR)–Gs protein complex [9].

GPCRs trigger the cyclic adenosine monophosphate (cAMP) signaling pathway used for cellular communication. Many components in the cAMP pathway have been associated to rafts [10]. There are also signs of activity–dependent partitioning of some GPCRs and G proteins into domains [11, 12]. GPCRs are characterized by rapid activation of signaling pathways, and the compartmentalization of signaling components to specific domains would enable the activation of the pathways more rapidly and with higher fidelity than in case all signaling components were randomly distributed along the membrane. Thus, it has been suggested that membrane domains would serve as a platform in the formation of GPCR-involving signaling complexes [13–15].

It remains unclear, however, whether the GPCRs reside in any specific membrane domain, since the studies investigating this topic have given contradictory results. GPCRs are widely associated with cholesterol [16–18], which implies that they would prefer raft-like domains enriched with cholesterol. On the contrary, some GPCRs have been found to prefer the L_d phase and the proximity of lipids they are comprised of [19–21]. The objective of this study is to find out the preference of β_2AR for different membrane domains. To this end, we studied the partitioning of β_2AR in a

highly phase separable model bilayer by MD simulations. β_2AR is one of the best characterized GPCRs with both its inactive [16] and active [9] crystal structures determined. It is expressed in pulmonary and cardiac myocyte tissue [12, 22] and important in the treatment of respiratory and heart diseases, such as asthma and hypertension [23, 24]. The knowledge of its function, regulation, and response is therefore of great importance.

Unfortunately, consideration of the phase segregation of lipids is currently challenging for atomistic MD simulations. This problem can be defeated by simplified coarse-grained models that enable studying of biomolecular processes that involve greater length scales and durations of tens of microseconds, which cover the lipid phase separation process. For the purpose of studying the partitioning of β_2AR in a phase separated bilayer, in this thesis a total of 22 coarse-grained systems were simulated for 10 μs each. To prevent initially placing the protein in a biased location in the membrane, the protein was let to adjust its location without control by initially placing it in different environments; in the L_o phase, in the L_d phase, and in a random lipid distribution. Additionally, two atomistic systems were simulated for 2 μs each in order to validate the results obtained in the coarse-grained simulations.

This thesis is divided into six chapters. The following chapter covers the biological background of the topics discussed in this thesis, such as lateral domains in membranes and the importance of GPCRs. The third chapter focuses on the methodological background describing the basics of molecular dynamics. The approach to the objective of this study is described in the fourth chapter, in which the simulation systems and the force field parameters employed are introduced along with the analysis methods. The results obtained from the simulations are presented and discussed in chapter five, and finally the conclusions drawn from this study are presented in chapter six.

2. BIOLOGICAL BACKGROUND

In this chapter the biological background of this thesis is presented. The structure of the cell membrane is briefly presented to start with. We now discuss how the concept of functional domains in cell membranes has modified the early fluid mosaic model of the cell membrane and is presented next, followed by a description of receptor proteins, which are important signaling molecules in membranes. Finally, a significant family of transmembrane receptor proteins, G protein–coupled receptors, is introduced along with some highlights of the β_2-adrenergic receptor, which is among the best-characterized GPCRs and the receptor used in the current study.

2.1 Structure of Cell Membrane

The cell membrane is crucial for the function of cells, since it separates the interior of the cell from the outside environment and thus protects the cell. In addition, the membrane prevents the contents from escaping and determines the molecules that can pass through it: it allows the nutrients to enter and the waste products to leave. When a cell is growing or changing, the membrane will change along it by creating a new membrane or by restructuring the existing one to make sure that there are no holes in the surface. It is vitally important for a cell to have a membrane since it could not survive or grow without it.

One of the earliest models of the structure of the cell membrane is based on the fluid mosaic model proposed by S.J. Singer and G.L. Nicolson in 1972 [25]. Many other membrane models have been proposed as well over time, but the general idea of the fluid mosaic model still remains as the best characterization of the structure and function of the plasma membrane. The fluid mosaic model states that the plasma membrane is a mosaic of different components, and that these components move laterally in the membrane making its character fluid-like in two dimensions [25]. The membrane components include different kinds of lipids and proteins that diffuse easily and serve important functions.

Over thousands of different kinds of lipids exist in cell membranes, such as glycerophospholipids, sphingolipids, and sterols [27]. These different types of lipids all share a key feature: they are dual-loving molecules [26]. Lipids have hydrophilic

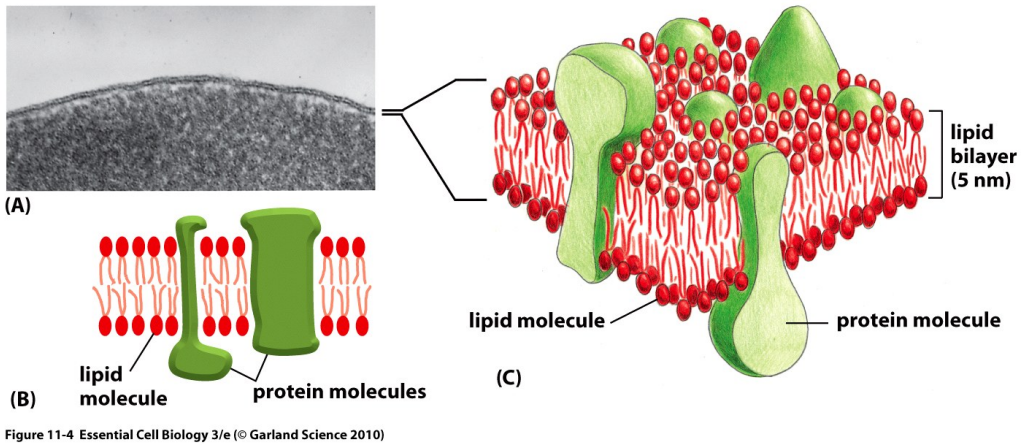


Figure 2.1: (A) An electron micrograph of human red blood cell plasma membrane, seen in cross section. (B and C) Schematic drawings of the lipid bilayer, lipids are shown in red and proteins in green [26].

head groups that prefer to stay in contact with an aqueous environment, and hydrophobic tails that repel water. This amphiphilic nature allows lipids to form the bilayer structure of a membrane, where the head groups face water from both sides of the membrane and tails stay inside the bilayer, away from water [26]. In addition to acting as membrane components, lipids are important, e.g., in signaling and energy storage [27].

Proteins form another major component in plasma membranes. They have important functions such as transporting nutrients across a bilayer, anchoring a membrane to macromolecules on either side, receiving and relaying signals, or working as enzymes to catalyze reactions [26]. Proteins consist of chains of amino acid residues whose sequence varies with different proteins [26]. Membrane proteins can be divided into integral membrane proteins, that are embedded to the bilayer, and to peripheral membrane proteins, that rest on the membrane surface by attaching to the bilayer or to an integral membrane protein [28]. In Fig. 2.1 the basic structure of the membrane is described by the fluid mosaic model: a lipid bilayer with embedded proteins. The mass ratio of proteins and lipids in the membrane composition can vary with cell type from ∼20:80 in myelin to ∼75:25 in mitochondrial inner membrane [28].

In addition to lipids and proteins, also carbohydrates are important components of plasma membranes. They prefer to stay on the outer leaflet of the plasma membrane, where they are bound to proteins forming glycoproteins or to lipids forming glycolipids [29]. Carbohydrates can form special sites on the membrane surface that

helps the cells recognize each other. This phenomenon is important for example for the immune system to function.

2.2 Lateral Domains Might Play Essential Roles in Membranes

The emphasis of the early fluid mosaic model was in the free lateral diffusion of lipids and proteins in the membrane, thus implying random order at least at long range, although the existence of small range order was not denied [25]. In 1982 Karnovsky et al. validated a theory of lipids organizing into lateral domains of functional and structural significance [30]. The formation of the domains was suggested to be driven by lateral phase separation of membrane components into liquid and gel phases, and cholesterol was found to play an important role in the domain formation [31, 32]. The concept of functional domains, called lipid rafts, was developed further proposing the rafts to be enriched with glycolipids, sphingolipids, and cholesterol present in cell membranes [33], and that the rafts serve as important platforms in the sorting of several membrane proteins and in the formation of signaling complexes [6].

The lipid rafts are currently considered as functional, small, dynamic, and ordered microdomains ranging from 10 to 200 nm in size, enriched with cholesterol and sphingolipids and having lifetimes in the range of microseconds to milliseconds [34–40]. The rafts are formed by the phase separation of liquid bilayers [41]. The lipid separation into two macroscopic phases, to a thicker, liquid-ordered (L_o) domain (raft–like) enriched with cholesterol, and to a thinner liquid-disordered (L_d) domain, has been found to occur in some synthetic bilayers composing of a ternary mixture of cholesterol, fully saturated lipids, and unsaturated lipids [42]. The phase separation is proposed to be highly cholesterol-dependent since cholesterol rings are planar and therefore prefer interaction with straight, saturated lipid chains over unsaturated lipid chains that are more bulky [43]. The membrane plane is thus physically segregated into the L_d phase, characterized by high mobility and disordered lipid chains, and to L_o phase, enriched with cholesterol that has an ordering effect on the lipid chains and thus slows down the lateral mobility and increases the thickness of the domain [44]. A schematic figure of the phases is shown in Fig. 2.2. The two liquid phases only coexist at certain temperature ranges and lipid ratios. A phase diagram with a highlighted phase coexistence region is shown in Fig. 2.3.

Caveolae are a subclass of lipid rafts containing a protein from the caveolin family, in addition to being enriched with cholesterol and sphingolipids [46]. The structure of caveolae was identified by electron microscopy in 1950s to be an ∼100 nm invagi-

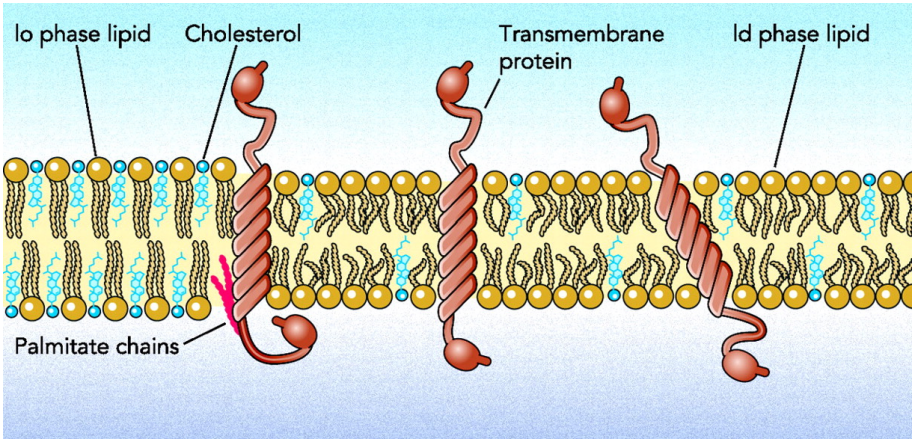


Figure 2.2: A shematic illustration of a lipid bilayer showing phase coexistence. The liquid-ordered domains are enriched with cholesterol (blue), being thicker and more ordered than the liquid-disordered domain. The transmembrane proteins locate to the membrane in such a way that the hydrophobic regions match. They might be targeted to the thicker L_o regions due to the hydrophobic matching or for example due to palmitoylation. The figure is modified from reference [45].

nation in the plasma membrane instead of being planar [47, 48]. Caveolin stabilizes the structure of caveolae [49, 50], and contains a scaffolding region that contributes to binding of the protein into a membrane. Caveolin proteins insert to phosphatidylcholine (PC) membranes in a manner that is highly cholesterol dependent [51]. Many membrane proteins participating in cell signaling have been found present in lipid rafts and caveolae, and therefore these regions are thought to regulate many kinds of cellular activities [52, 53], such as endocytosis, exocytosis, transcytosis, nutrient transport, viral entry and budding, and also receptor and ion channel expression, activation and desensitization [54–58]. According to the caveolin signaling hypothesis, the interaction of signaling proteins with the caveolin scaffolding domain alters the signal transduction partners by inhibiting the basal state signaling and facilitating the interaction of components upon activation of the signaling pathway due to co-localization in caveolae [10].

The concept of lipid rafts has been controversially argued [60–64], since rafts are experimentally difficult to observe unambigously in vivo with current techniques. Rafts have been widely studied by isolating them as detergent-resistant membranes (DRMs) from cells, and these DRMs were thought to reflect the actual lipid-protein assemblies in living cells [6]. This resulted in studying and categorization of membrane proteins into raft and non-raft proteins, based on their detergent-solubility [65], however it was found that the group of proteins isolated in DRMs

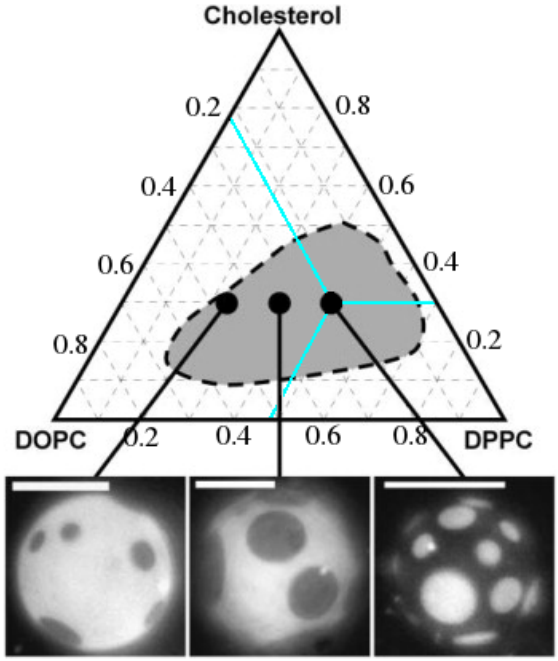


Figure 2.3: A phase diagram of a micron–scale liquid immiscibility region in GUVs. The liquid–liquid coexistence region at 25 °C is shown in gray inside the triangle–graph. Below the phase diagram are the fluorecence micrographs of GUVs studied by nuclear magnetic resonance (NMR), that show phase coexistence with different DOPC:DPPC lipid ratios: 2:1 (left), 1:1 (center) and 1:2 (right), all mixed with 30 mol-% of cholesterol. The scale bars are 20 μm. The graph is read as follows: the right face of the graph shows the mass fraction of cholesterol, left face shows the mass fraction of DOPC and bottom face indicates the mass fraction of DPPC in the membrane. For the ratio of DOPC:DPPC:CHOL being 0.23:0.47:0.30 in the left, lines can be drawn from those points in the sides towards the center of the triangle (shown in cyan). The cross point of the lines shows the phase of the lipid mixture, for example in this case two liquid phases are present. Other possible phases outside this region are not shown in the graph. The figure is modified from reference [59].

depended on the detergent and conditions used [66]. It was also shown that some detergents used could grow or even create ordered domains in an initially homogeneous membrane [67], and thus DRM experiments could not be considered valid for the prediction of protein or lipid composition in the assumed domains in live cell membranes [62].

The lipid phase separation was found to occur in synthetic model systems, so called giant unilamillar vesicles (GUVs), yielding more pronounced phase separation, the more pronounced differences there were between the properties of the lipids [3]. In Fig. 2.3 a confocal fluorecence microscopy figure of the phase coexistence on a GUV is shown. There is evidence that the phases could coexist in cell membranes as

well [68–70]. The plasma membrane of a live cell does indeed contain lipids necessary for the phase separation to occur. However, the conditions in live cell membranes are much more complex than in synthetic model membranes. Live plasma membranes consist of many different lipid species [71] than the three types needed for the L_o/L_d phase coexistence. In addition, a major amount of the membrane area is occupied by proteins that interact with the lipids [72]. The proteins have even been found to destabilize the phase separation [73]. Also the human body temperature sets limits to the phase co-existence, at least optically resolvable domains are rare at temperatures that high [74]. As a consequence to this complexity, the existence of rafts in synthetic membranes does not necessarily signify their existence in live cells, despite the evidence of existing L_o and L_d domains and proteins prefering one domain over the other. On the other hand, the assumption of rafts not existing is not more justified either.

2.3 Receptors Mediate Signals and Interact with Lipids in the Membrane

Membrane receptors can be found on the surface of the cell, mediating information from the exterior of the cell to the interior of it by triggering a response upon activation. Depending on type, receptors' information transfer can be triggered by binding of signaling molecules, called ligands, or they can respond to light, odors, or change in the oxygen concentration [28]. A given ligand can trigger different responses depending on the receptor to which it binds. On the other hand, receptors are not ligand-specific, since many ligands can bind to a certain receptor [75]. Different receptor families exist, such as ion channel receptors, enzyme-linked receptors, and G protein-coupled receptors [26]. Cells contain multiple receptors of different kinds depending on cell type.

Transmembrane receptors have an extracellular domain receiving the signal, a hydrophobic region spanning the membrane, and an intracellular domain relaying the signal. Once a receptor receives a signal, it activates, and the receptor goes through a conformational change that might enable the intracellular domain to bind an effector protein or open an ion channel. Most of the existing drugs are targeted to receptors to modify their activity, either by activating the receptor, or by preventing its activation in order to obtain the desired response.

Many membrane proteins have been shown to require specific lipids for stabilization and for activity [76–78]. It has been suggested that the variation of the composition of a lipid bilayer might affect the function and structure of membrane

proteins [79]. Lipids might affect the protein function either by interacting directly with the protein, or via indirect interaction by changing the membrane physical properties [80, 81]. Hydrophobic or hydrophilic lipid-protein interactions have been found to result in a slower lipid exchange near the protein surface, and therefore proteins are surrounded by a shell of annular lipids that behave differently than the lipids further from the protein [82]. The movement and order parameter of these boundary lipids might be affected up to 3-4 nm away from the protein surface [83]. It has been estimated that 50-100 lipid molecules diffuse as a dynamic complex with the protein they are surrounding [83].

The preference of proteins for specific lipid domains can be tightly regulated by several mechanisms. In order to form a stable membrane structure, transmembrane proteins are adjusted to a membrane by hydrophobic matching, which means that the hydrophobic thickness of the protein matches with the hydrophobic thickness of the membrane [84–87]. If the hydrophobic regions are mismatched, the lipids near the membrane protein will adapt to the hydrophobic part of the protein by elastic distortion to minimize energy [84, 88]. The concept of hydrophobic matching has been proposed to be important in the formation of lipid rafts and for the selective partitioning of transmembrane proteins into specific lipid domains, where hydrophobic thickness of the domain matches with the hydrophobic part of the protein [85, 89]. Many receptors have been found present in lipid rafts, such as some epidermal growth factor receptors, platelet-derived growth factor receptors, and endothelin receptors [38]. Thus the signaling could be altered by the changes in the partitioning of molecules between raft and nonraft domains in the membrane. Also palmitoylation, covalent attachment of fatty acids to some residues of membrane proteins, has been reported to regulate the affinity of integral membrane proteins for rafts [90]. Some peripheral membrane proteins, such as a nonreceptor tyrosine kinase, has been found to be reversibly palmitoylated, losing the raft association after depalmitoylation [91]. It was also suggested that the affinity of proteins for certain domains is defined by lipid shells surrounding them, causing the protein-lipid complex to target domains consisting of lipids compatible with the lipids of the shell [82].

2.4 G Protein-Coupled Receptors – An Important Receptor Superfamily

The family of GPCRs consists of around 800 members [7] and makes up the largest family of membrane proteins. GPCRs are involved in a wide range of physiological processes related to several diseases, i.e., metabolic [92–94], mental [95,96], immunological [97,98], inflammatory [99], cardiovascular [97,100], and senses [101] disorders, as well as cancer [102]. The variety of physiological processes involving GPCRs has made them an attractive target for the development of drugs, approximately 30-50 % of the current drugs acting on GPCRs [103, 104], yet on only a small fraction of them [105].

GPCRs share a charasteristic structure consisting of seven transmembrane α-helices connected to each other with three extracellular and three intracellular loops, and an eighth short helix parallel to the membrane intracellular surface. The amino acid chain starts from an N-terminus outside the cell and ends with a C-terminus inside it. The extracellular domain of the receptor acts as a ligand binding site whereas the intracellular domain binds the guanine nucleotide-binding protein (G protein). Despite the common core in all GPCRs, the intracellular and extracellular regions are different in structure, length, and sequence, enabling then to respond to a wide variety of ligands, such as ions, amines, odorants, lipids, peptides, proteins, nucleotides, and photons [106]. The first classification method of GPCRs divided them into six classes from A to F based on the sequence homology: Class A (rhodopsin-like), B (secretin receptor family), C (metabotropic glutamate), D (fungal mating pheromone receptors), E (cyclic AMP receptors) and F (frizzled or smoothened) [107]. An alternative classification of GPCRs has been suggested, called GRAFS, and it divides GPCRs into five main families according to their evolutionary origin: glutamite, rhodopsin, adhesion, frizzled, and secretin, of which rhodopsin forms the largest group [108].

Once a GPCR is activated, it binds specific G proteins, of which most are heterotrimers consisting of Gα, Gβ, and Gγ subunits. The Gα and Gγ subunits are anchored to the membrane through a lipid chain. The Gα subunit consists of an α-helical domain and of a Ras domain, which is a GTPase, meaning that it can bind and hydrolyze guoanosine triphosphate (GTP). In an inactive state the subunit α binds guanosine diphosphate (GDP), and activates once it is replaced by GTP. Activation proceeds as follows: binding of ligand induces a conformational change in the receptor, which further on enables the binding of G protein in the intracellular

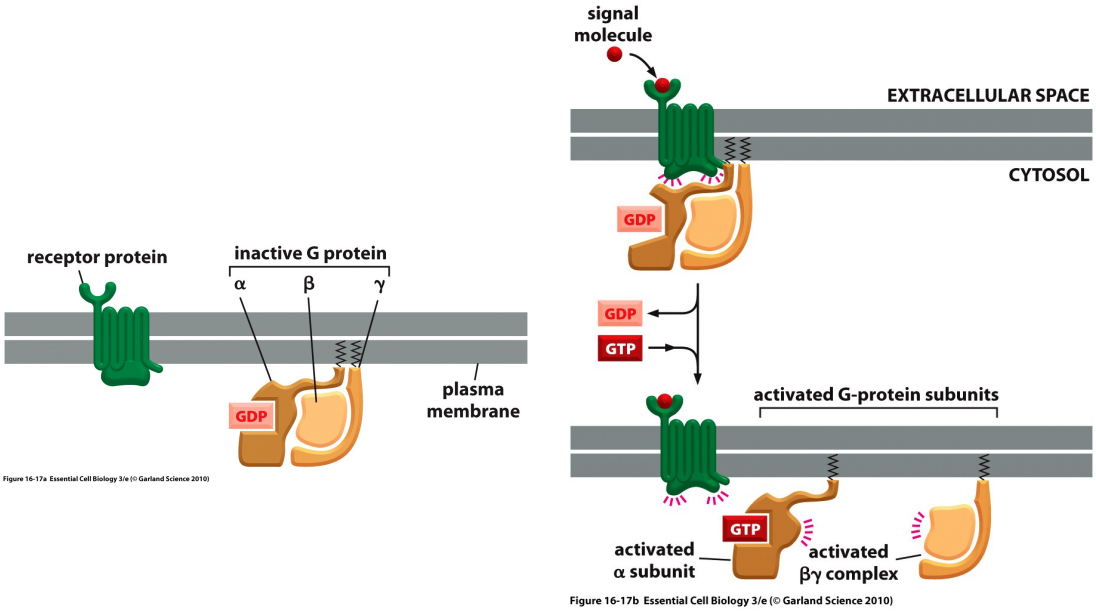


Figure 2.4: The activation process of GPCRs. (a) The inactive state of both the receptor and the G protein. (b) The binding of ligand induces conformational changes in the receptor, that enables the binding of the G protein. The Gα subunit is able to exchange its GDP for GTP. As a result, both the α subunit as well as the $\beta\gamma$ complex are activated and able to interact with their targets in the membrane. After the subunits have transmitted the signal and the ligand is no longer bound to the receptor, the components return to the inactive state, shown in (a). The figures are taken from reference [26].

domain, allowing the Gα subunit to release the bound GDP and bind GTP instead. The Gα-GTP complex decouples along with G$\beta\gamma$ to interact with a downstream effector protein, after which the GTP is hydrolyzed. The G protein then reforms back to trimer, ready for a new activation process. The described process is demonstrated in Fig. 2.4. GPCRs trigger the cAMP signaling pathway, in which the different G proteins, Gs and Gi, cause opposite reactions. An activated Gsα subunit activates an enzyme called adenylyl cyclase, which in turn converts adenosin triphosphate (ATP) into cAMP. In contrast, the Gi protein reduces cAMP levels and inhibits adenylyl cyclase. Increased cAMP will contribute to relaxation in smooth muscle, together with increasing contractility and pulse rate in cardiac muscle [109]. The correct operation of the cAMP pathways is crucial, since a deregulation of the cAMP pathways and an aberrant activation of cAMP-controlled genes have been associated to the growth of some cancers [110, 111].

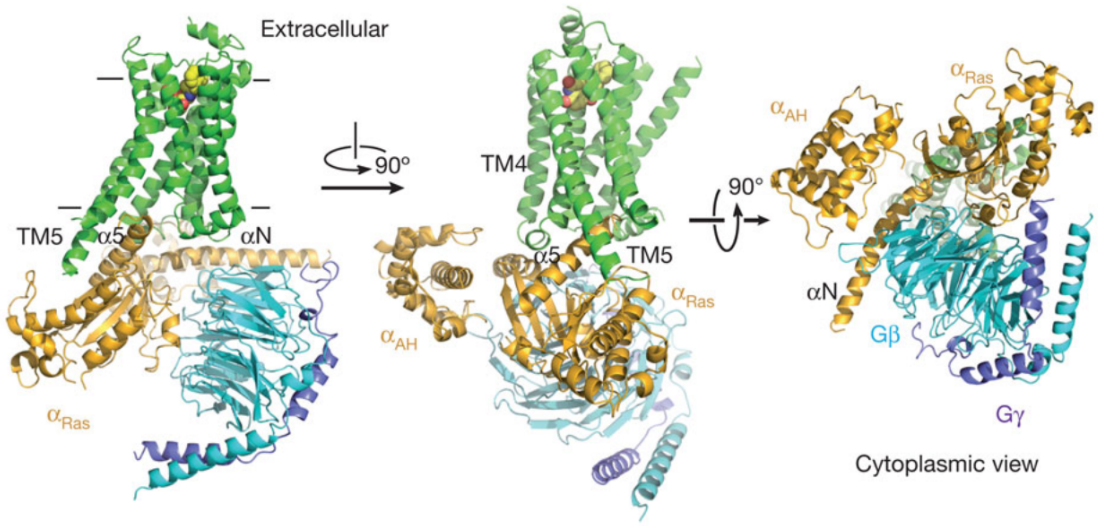


Figure 2.5: The structure of the β_2AR-Gs complex. The receptor is shown in green with a ligand bound to the extracellular pocket, Gα subunit is bound to the receptor's intracellular domain and shown in gold, Gβ subunit is shown in cyan, and Gγ subunit in violet. The figure is taken from reference [9].

2.4.1 Structure–Function Relationship of GPCRs

Several X-ray structures have been solved for rhodopsin-like GPCRs during the past years, the first one being bovine rhodopsin, later accompanied by β_1- and β_2-adrenergic receptors, dopamine D3 receptor, adenosine A2A receptor, histamine H1 receptor, muscarinic M2 and M3 receptor, etc. [112]. The importance of the work within the structure discovery is highlighted by the Nobel Prize awarded in 2012 for the discovery of the X-ray structure of β_2-adrenergic receptor–Gs protein complex (shown in Fig. 2.5) [9]. The discovered structures will enable further research on more detailed properties of the receptor, such as function, activation, and regulation.

The residues in the cavity inside GPCRs at the extracellular side have been shown to interact with ligands, triggering an outward tilting of the extracellular end of the transmembrane helix 6 [113]. The movement causes conformational changes that allow G protein coupling [114, 115]. These conformational changes occurring upon activation are thought to be general across all rhodopsin-like GPCRs [116]. The two states, activated and inactivated, are both assumed to exist in vivo as an equilibrium controlled by some amino acids acting as micro-switches [117–120]. Agonists and antagonists may stabilize either state by interacting with these residues. This mechanism is regulated by different amino acids depending on receptor, even though the mechanism is general [121]. X-ray structures involving bound agonists have been

found for β_1- and β_2-adrenergic receptors as well as for adenosine A2A receptor during the recent years, and more are expected to come in the future [112]. The structures will provide new insights for the research and for the rational drug design of agonists and antagonists, elucidating the advantage of modeling even broader structural changes [114, 122]. Any knowledge available on the receptor function and on the factors modulating it can be important in this process.

2.4.2 β_2-Adrenergic Receptor Is a Characteristic GPCR

Adrenergic receptors (AR) are well-characterized GPCRs belonging to the rhodopsin family, and responsible for example for triggering increased heart rate and mobilization of energy in the so called fight-or-flight response [28]. They mediate signals from the sympathetic nervous system to the cardiovascular system [123], responding to ligands such as adrenaline in bronchial vasculature and noradrenaline in cardiac muscle. The subtypes of adrenergic receptors, α and β, have different locations and effects, and bind to different G proteins. Gs protein, which activates adenylyl cyclase and causes smooth muscle relaxation, heart muscle contraction, and glycogen break down, is bound by β_1- ,β_2- and β_3-adrenergic receptors. Since β-adrenergic receptors are highly important in cardiovascular and pulmonary physiology, they have many agonists and antagonists with medical applications. β_2AR agonists exist that are asthma drugs, causing smooth muscle relaxation, rapid release of insulin, bronchial passage dilation, and vasodilation, mimicking the effect of native ligands [24]. Some βAR antagonists are used as beta blockers treating heart disease and hypertension, slowing the heart rhythm and reducing blood pressure, hence diminishing the effects of adrenaline [23]. This Thesis focuses on the properties of β_2AR.

β_2AR is expressed in pulmonary and myocyte tissue [12, 22]. It has been found to couple to both Gs and Gi proteins in cardiac myocytes [124], but also to signal without G protein, via arrestin [125, 126]. The structure of β_2AR consists of seven transmembrane helices typical for GPCRs, including a short eighth helix. Five of the helices have a kink in the middle of the receptor, producing a structure pinched in at the waist with cavities on both sides of the membrane. The structure of β_2AR in an inactive state has been determined with bound inverse agonists timolol [18] and carazolol [16]. Also the active structure of β_2AR has been discovered bound to an agonist and to Gs protein [9], shown in Fig. 2.5. The activation of β_2AR involves a 7-14 Å outward movement of the intracellular end of the helix 6, and a 2 Å inward movement of the extracellular end of the helix 5 [127, 128], shown in Fig. 2.6. The Ras domain of the Gα subunit is in contact with the transmembrane helices 5 and

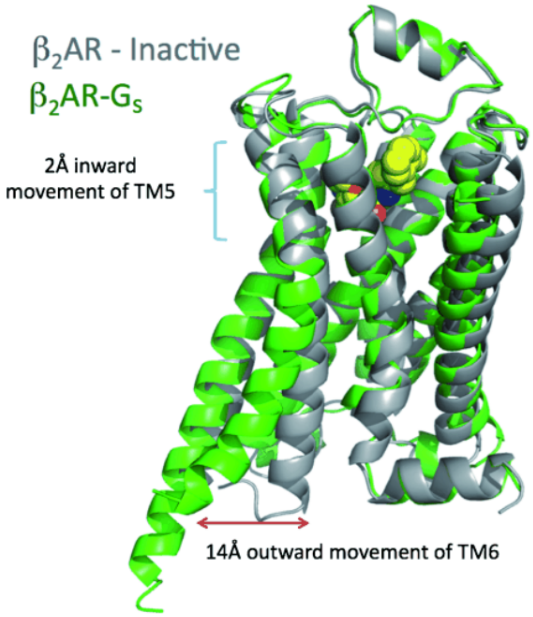


Figure 2.6: The structure of β_2AR in an inactive state (gray) compared to the active state (green) of the β2AR-Gs complex. The figure is taken from reference [127].

6 of β_2AR via its C-terminal helix, and this binding site exists only when β_2AR is in its active form. Cholesterol has been found to increase the kinetic, energetic, and mechanical stability of β_2AR [17], and is also required for its crystallization [16,18]. In addition to cholesterol, also phospholipids have recently been found to regulate the activity of β_2AR [129].

2.4.3 GPCRs Have Functional Consequences with Domains

It was proposed earlier that the GPCRs and G proteins are homogenously distributed along the membrane, diffusing freely and interacting in case they collide [130,131]. However, this perspective has been questioned since the abundance of GPCRs, G proteins, and their effector proteins is overall too low to provide sufficient enrichment for rapid signaling processes that are characteristic for GPCR activation [132]. The concept of special membrane domains acting as compartmentalized signaling complexes has provided a new perspective for GPCR signaling [13–15].

Several GPCRs have been found to localize in rafts or caveolae, either before or after treatment with agonists [10]. For some GPCRs, agonists can promote entering into rafts or caveolae, whereas for others, agonists have an influence promoting the partitioning of the receptors out from these domains, which has been suggested to

initiate receptor internalization or endocytosis and desensitization [54,58]. Adenylyl cyclase, cyclic nucleotide phosphodiesterase, and protein kinase A, important signaling components in the GPCR triggered cAMP pathway, can target to rafts or caveolae [10]. Heterotrimeric G proteins have been found to interact directly with caveolin proteins and to localize to caveolae [132, 133]. The interaction with caveolin has been found to help maintain the Gα subunits in an inactive state [134,135], and caveolin scaffolding region is suggested to regulate the G protein function [135]. It has been later suggested that heterotrimeric G proteins dynamically partition to lipid rafts surrounding GPCRs more preferentially than monomeric Gαq [12]. This signifies activity-dependent partitioning of G proteins that was suggested to be driven by their lipid anchors [12]. Organization of the signaling components can compartmentalize cAMP signals by creating cellular microdomains with either higher or lower levels of cAMP [10]. The localization of the components to specific regions enables agonists to activate signaling pathways rapidly and with high fidelity, especially when the signaling pathways require interaction of multiple components [10].

As integral membrane proteins, a significant portion of GPCRs is embedded in the membrane. For rhodopsin, the lipid–protein interface has been estimated to correspond to $\sim$38% of the total receptor surface area [136]. Due to the large contact area with the membrane, the lipid environment is thought to have an important effect on the receptor structure and function [76]. It has been proposed that targeting GPCRs to lipid rafts and caveolae is one possible mechanism by which the membrane alters the GPCR function [132,137,138]. One factor suggested to address the GPCRs to ordered lipid rafts or caveolae domains is fatty acid modifications, such as palmitoylation of cysteine residues in the C-terminal helix [139]. Palmitoylation has been found to influence receptor activity, trafficking, and G protein coupling for at least serotonin and oxytocin receptors [140]. Membrane cholesterol has been shown to modulate the function of many GPCRs, although the mechanism is not fully understood [141]. It has been suggested that cholesterol modulates the GPCR function either directly inducing conformational change [142, 143], or indirectly by altering the membrane physical properties [76, 144]. In addition to cholesterol, also phospholipids have been found to regulate allosterically GPCR activity [129]. The type of lipids directly interacting with GPCRs, such as cholesterol and certain phospholipids, might also define the affinity for domains consisting of lipids compatible with the lipids surrounding the GPCR [82].

2.4.4 Current State in the Research of Membrane Domains and GPCR Partitioning

The phase separation of lipids has been observed in model membranes of synthetic lipids [3–5], where more pronounced differences in the properties of lipids yielded more pronounced phase separation [3]. Certain proteins and lipids have been found to be detergent insoluble [3], and proteins have been found to phase partition in giant plasma membrane vesicles [74] and in GUVs [19]. The phase separation of lipids has also been shown computationally by MD simulation studies. In 2008 Risselada and Marrink carried out simulations of a lipid bilayer comprised of saturated phospholipids, polyunsaturated phospholipids, and cholesterol, obtaining spontaneous separation of lipids into two phases [145]. One of the two phases was mainly composed of saturated lipids and cholesterol, and the other contained only unsaturated lipids and less cholesterol. The compositions of the two domains were found to be consistent with a study based on NMR measurements having a similarly phase separated model membrane [59]. Schäfer et al. observed in their joined study of computational biology and confocal fluoresence microscopy of GUVs that model transmembrane helices, such as WALP peptides, spontaneously partition into the L_d phase [146], which was also found in another computational study by Domanski et al., along with the result that the rhodopsin receptor as well prefers the L_d phase [20]. A study by de Jong et al. interestingly showed that palmitoylation of WALP peptides prodvided a driving force towards the more ordered membrane domains, although the two saturated lipid anchors in the study were not sufficient to allow the peptides to totally dissolve into the L_o domains [147].

It remains still ambiguous where the GPCRs reside in a membrane that is segregated into two lipid phases, since several studies give contradictory results. For many GPCRs the preference for domains has been proposed based on experimental findings [10], for example βARs. Findings from some DRMs studies in 2001 suggested rhodopsin to locate in the L_o domains [148], although it was shown by a GUV–based study in 2005 [19] to prefer the L_d phase (Fig. 2.7). This result was repeated by simulations in 2012, where rhodopsin was found to prefer the L_d phase irrespective of hydrophobic mismatch, and it was suggested that the enthalpically unfavorable disturbance of the tightly packed lipids in the L_o domains drives the protein into the L_d domains [20]. Experimental study of rat ventricular cardiomyocytes from the year 2000 shows localization of β_1- and β_2-adrenergic receptors to caveolar domains where β_2AR, but not β_1AR, was found to translocate out of cardiomyocyte caveolae

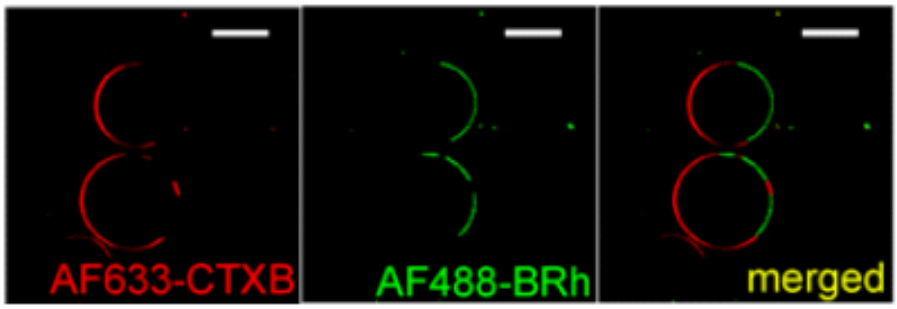


Figure 2.7: A dual-color confocal microscopy image of a bacteriorhodopsin–containing giant unilamellar vesicle. AF633-CTXB (red) and AF488 (green) are fluorecent labels, of which AF633-CTXB (right) reveals the region of the L_o phase, whereas AF488 (center) shows the L_d phase. AF488 was distributed to membrane areas complementary to those in which AF633-CTXB was localized, as shown in left where the two first figures are merged. Bacteriorhodopsin is labeled with AF488, and thus shows preferential partitioning into the L_d phase. The figures are taken from reference [19].

upon agonist stimulation [11]. On the other hand, since cholesterol is widely associated with β_2AR and needed for its stabilization [16–18], β_2AR might be assumed to prefer the L_o phase enriched with cholesterol. As it was postulated, the type of lipids surrounding the protein might target the protein and the lipids surrounding it to certain domains. However, there is no clear view where β_2AR is located. The objective of this study is to find out the β_2AR preference for the L_o and the L_d domains in a phase separated lipid bilayer by a MD simulation approach.

3. MOLECULAR DYNAMICS

Biological processes occur over a wide range of spatial and time scales, for which reason the method of the study should be chosen carefully. Whenever the processes involve time scales too fast for human eye or length scales beyond microscopic limits, computational simulations are an attractive alternative. They can provide details that are not accessible through experiments, such as individual particle motions in time, and provide access to unraveling more specific questions.

Molecular dynamics simulations have become a standard tool in the investigation of molecular level phenomena of biomolecules [149]. To date, it has yielded success for example in the detailed investigation of processes occurring in cell membrane level, such as receptor–ligand and receptor–lipid interactions related to drug development [150]. MD simulations are based on classical physics, mimicking the motions and distributions of atoms and molecules in real life. In terms of time and length scales, simulations can now reach into the microscopic regime, forming a bridge to experiments such as vesicle aspiration, fluorescence imaging, and atomic force microscopy measurements [151]. Simulations are indeed often used to interpret experiments, but also to discover new properties beyond current experimental techniques [152]. The field of MD simulations is developing rapidly and becoming more and more efficient, along with the development of computer hardware. There are several software packages available for MD simulations of biological systems, such as GROMACS, NAMD, and CHARMM. The simulations in this study were carried out with the GROMACS software package [153].

The basics of the MD method is presented in the following sections. Force field defines the interactions between the simulated particles, the time evolution is based on the numerical evaluation of Newtonian equations, thermostat and barostat algorithms regulate the system properties, and boundary conditions ensure that the system does not have an unrealistic border with the vacuum. Next, the simulation process, formed from the previous elements, is described, after which also the limitations of MD are discussed. The studied system could be represented at different levels of detail. Atomistic representation best reproduces the real-world systems, yet coarse-grained representations are increasingly popular when large systems or long

time scales are required [154]. In this study of β_2AR partitioning in a phase segregated membrane the Martini force field for coarse-grained molecules was used since the phase separation and protein partitioning involve spatial and time scales that are challenging for atomistic simulations. The coarse-grained model is presented briefly in the end of this chapter.

3.1 Force Field Describes the Interactions

The particles simulated by MD obey the laws of classical mechanics in their motion, that is caused by interactions with the other simulated particles. The interactions are defined by the potential equations of the chosen force field. There are different kinds of force fields in use for varying purposes; reactive force fields taking into account chemical reactions [155], polarizable force fields focused on electronic polarizability [156], all-atom force fields including specific representation of all involved atoms [157], coarse-grained force fields using simplified models of functional groups [158], etc. The choice of the force field must take into account the system properties as well as the research question in order it to best suit the studied system and correspond to the forces experienced between real particles. The force fields used in this thesis were the coarse-grained Martini 2.2 force field and the all-atom OPLS (Optimized Potentials for Liquid Simulations) force field. The particles in the force field are considered as moving mass points, and the time evaluation of their movement is obtained by numerically integrating the Newton's equation of motion. Two types of interactions can take place between the particles: bonded interactions between defined groups of particles, and non-bonded interactions between particles, based on their changing distance [159].

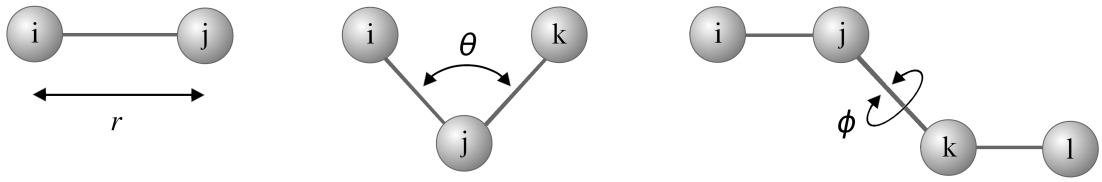
3.1.1 Bonded Interactions

Bonded interactions concern stretching, angle vibration, and dihedral angle torsion of covalent bonds. The covalent bond stretching between two bonded atoms i and j, illustrated in Fig. 3.1(a), can be described by a harmonic potential of the form [159]

$$V_{\text{bond}}(\mathbf{r}_i, \mathbf{r}_j) = \sum_{\text{bonds}} \frac{1}{2} k_r (r - r_0)^2, \tag{3.1}$$

where

$$r = |\mathbf{r}_i - \mathbf{r}_j| \tag{3.2}$$



(a) (b) (c)

Figure 3.1: Bonded interactions. The principle of (a) bond stretching, (b) bond angle vibration, and (c) proper dihedral torsion.

is the distance between the atoms, k_r is the harmonic force constant, and r_0 is the reference distance between the two atoms i and j.

The vibration of a covalent bond angle formed between three atoms i, j, and k, shown in Fig. 3.1(b), is described by a harmonic angular potential given by [159]

$$V_{\text{angle}}(\mathbf{r}_i, \mathbf{r}_j, \mathbf{r}_k) = \sum_{\text{angles}} \frac{1}{2} k_\theta (\theta - \theta_0)^2, \tag{3.3}$$

where

$$\theta = \arccos \frac{\mathbf{r}_{ij} \cdot \mathbf{r}_{kj}}{r_{ij} r_{jk}} \tag{3.4}$$

is the angle between the bonds, and θ_0 is the reference angle.

The dihedral angle ϕ is defined by the positions of atoms i, j, k, and l as the angle between the normals to the two planes (i, j, k) and (j, k, l) (Fig. 3.1(c)) [159]. An angle of zero corresponds to the *cis* and an angle of 180^o corresponds to the *trans* configuration. The proper dihedral potential is described by a periodic function given by [159]

$$V_{dihedral}(\phi) = k_\phi (1 + \cos(n\phi - \phi_0)), \tag{3.5}$$

where k_ϕ is the force constant, n the periodicity, and ϕ_0 the reference angle. Instead of using an interaction between atoms i and l, a set of periodic functions with different periodicity can be used, or a set of powers of cosine functions as in the Ryckaert-Belleman's (RB) potential given by [153]

$$V_{\text{RB}} = \sum_{n=0}^{5} C_n cos^n \phi, \tag{3.6}$$

where $C_n (n = 0, ..., 5)$ are the Ryckaert-Belleman torsion parameters. GROMACS

uses the Ryckaert-Bellemans equation to compute Fourier dihedrals of the form [153]

$$V_{\mathrm{F}}(\phi_{ijkl}) = \frac{1}{2}[C_1(1+\cos(\phi)) + C_2(1-\cos(2\phi)) + C_3(1+\cos(3\phi)) + C_4(1+\cos(4\phi))]. \tag{3.7}$$

Additionally, improper dihedrals are defined in order to prevent molecules from flipping over to their mirror images or to keep planar groups planar, such as aromatic rings. The improper dihedrals are based on four atoms i, j, k, l as the proper dihedrals, and given a harmonic restraining potential [159]

$$V_{\mathrm{improper}} = \frac{1}{2}k_\xi(\xi - \xi_0)^2, \tag{3.8}$$

where k_ξ is the force constant, ξ the improper dihedral angle between the planes (i, j, k) and (j, k, l), and ξ_0 is the corresponding referencem angle.

3.1.2 Non-bonded Interactions

Non-bonded interactions act between particles that are not linked to each other by a covalent bond. They are defined as a function of the distance r_{ij} between the two particles. The non-bonded interactions are usually considered with a certain cut-off radius in the simulated system, meaning that after excluding pairs that are already involved in the bonded interactions, the non-bonded interactions are only calculated for pairs whose distance r_{ij} is less than the cut-off radius. The particles involved in the non-bonded interactions will thus fluctuate during the simulation and must be updated regularly. The pairs for which the non-bonded interactions are calculated, are added to a neighbour list [153]. The non-bonded interactions are formed from the sum of van der Waals interactions and Coulomb interactions, of which the van der Waals potential has the form of the Lennard-Jones (LJ) potential given by [157]

$$V_{\mathrm{LJ}}(r_{ij}) = 4\epsilon_{ij}\left(\left(\frac{\sigma_{ij}}{r_{ij}}\right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}}\right)^{6}\right), \tag{3.9}$$

where σ_{ij} and ϵ_{ij} are constants giving the reference distance and minimum energy of the potential well. The first term of the equation describes a strong short-range repulsion, taking place if the particles are overlapping. The second term describes a longer-range dispersion between the two particles.

Coulomb interactions take place between charges or partial charges of atoms and

is given by [157]

$$V_{\mathrm{C}}(r) = \sum_{i<j} \frac{q_i q_j e^2}{\epsilon_{ij} r_{ij}} f_{ij}, \tag{3.10}$$

where e is the elementary charge and q_i, q_j are the charges of particles i, j, respectively. The total potential acting on a particle can be calculated as a sum of all the potentials [153]:

$$V_{\mathrm{tot}} = \sum V_{\mathrm{bonded}} + \sum V_{\mathrm{non-bonded}}. \tag{3.11}$$

The Coulomb potentials do not decrease rapidly with distance like the van der Waals interactions, implying that the electrostatic interactions beyond the cut-off radius can not be completely neglected. Methods such as the Ewald summation are commonly used for efficient computational evaluation of these interactions [160].

In addition to the potentials presented above, specially defined potentials can be used for restricting the motion of the system in order to for example avoid rearrangement of certain parts during the equilibration of the system or to include some knowledge obtained from experimental studies. These position restraints are used to keep particles in a fixed reference position by imposing forces on the given particles [153].

3.2 Time Evolution Stems from Solving the Equations of Motion

The time evolution of the system in MD simulation is obtained by numerically solving the Newton's equation of motion. For a system of N interacting atoms the Newton's equation of motion are given by

$$m_i \frac{\partial^2 \mathbf{r}_i}{\partial t^2} = \mathbf{F}_i, \qquad i = 1...N, \tag{3.12}$$

where m_i is the mass of the particle i, $\mathbf{r}_i$ is the position of the particle at time t, and $\mathbf{F}_i$ is the force acting on the particle. The forces can be obtained from the negative derivative of the potential function $V(\mathbf{r}_1, \mathbf{r}_2, ..., \mathbf{r}_N)$:

$$\mathbf{F}_i = -\frac{\partial V}{\partial \mathbf{r}_i}. \tag{3.13}$$

The equations (3.12) and (3.13) are computed using a time integration algorithm. GROMACS uses the so-called leap-frog algorithm [161] as default for the integration of the equation of motion [153]. It updates the positions and the velocities in different

points of time using the forces $\mathbf{F}(t)$ determined by the positions of the particle at time t. The equation for the velocity in the leap-frog algorithm is given by

$$\mathbf{v}(t + \frac{1}{2}\Delta t) = \mathbf{v}(t - \frac{1}{2}\Delta t) + \frac{\Delta t}{m}\mathbf{F}(t), \tag{3.14}$$

and the equation for the position is given by

$$\mathbf{r}(t + \Delta t) = \mathbf{r}(t) + \Delta t(t + \frac{1}{2}\Delta t), \tag{3.15}$$

where $\mathbf{v}$ is the velocity at time $t - \frac{1}{2}\Delta t$ and $\mathbf{r}$ is the position at time t [153].

3.3 Temperature and Pressure Coupling Maintain the Correct Ensemble

It is almost always necessary to implement some modifications to the Newtonian equations in order to avoid undesirable effects resulting from inexact solutions of the equations. Usually it is desirable to simulate the system in constant temperature and pressure, in a so-called isobaric-isothermal NpT ensemble, as experiments are often carried out in similar constant conditions. GROMACS has algorithms called thermostats and barostats designed for the temperature and pressure coupling, respectively; with these algorithms the temperature and the pressure of the system can be controlled during the simulation [153].

For simulation in constant temperature, GROMACS can use several thermostats such as Berendsen, velocity rescale, Andersen, and Nosé-Hoover [153]. The weak-coupling scheme of Berendsen uses an external heat path with temperature T_0. The algorithm slowly corrects the deviation of the system temperature T to the temperature of the heat bath T_0 by following the equation

$$\frac{\mathrm{d}T}{\mathrm{d}t} = \frac{T_0 - T}{\tau}, \tag{3.16}$$

where τ is the time constant for exponential decay [153]. The Berendsen method is efficient in the equilibration phase of the system, but it does not produce a correct ensemble since it suppresses the fluctuations of the kinetic energy and pressure. The velocity rescaling method is an improved version of Berendsen thermostat that creates a correct kinetic energy distribution due to an additional stochastic term,

resulting inthe form:

$$\mathrm{d}K = (K - K_0)\frac{\mathrm{d}t}{\tau_T} + 2\sqrt{\frac{KK_0}{N_f}}\frac{dW}{\sqrt{\tau_T}}, \tag{3.17}$$

where K is the kinetic energy and K_0 the kinetic energy at equilibrium, dW a Wiener process, N_f is the number of degrees of freedom, and τ_T is close to the time constant τ defined by

$$\tau_T = \frac{N_{df}k\tau}{2C_V}, \tag{3.18}$$

where N_{df} is the total number of the degrees of freedom, k is the Boltzmann's constant, and C_V is the total heat capacity of the system [153].

The system pressure coupling can be done in GROMACS for example by Berendsen, Parrinello-Rahman, velocity Verlet, or MTTK (Martyna-Tuckerman-Tobias-Klein) methods [153]. The Parrinello-Rahman algorithm changes the periodic box size in order to force the current pressure $\mathbf{P}$ towards the desired reference pressure $\mathbf{P}_{ref}$. Both $\mathbf{P}$ and $\mathbf{P_{ref}}$ are represented as tensors, each element of which is the force that acts on the surface of an infinitesimal cubic volume. In isotropic situations, where forces are same in all directions and in the absence of viscous force, the pressure tensor is diagonal and usually referred to as a scalar quantity. In material science studies, however, the pressure is referred to as a stress tensor, the diagonal elements of which are known as the tensile stress and the non-diagonal elements as the shear stress. The box vectors are represented by a matrix $\mathbf{b}$, and they are forced to obey the matrix equation of motion given by

$$\frac{\mathrm{d}\mathbf{b}^2}{\mathrm{d}t^2} = V\mathbf{W}^{-1}\mathbf{b}'^{-1}(\mathbf{P} - \mathbf{P}_{ref}), \tag{3.19}$$

where t is time and $\mathbf{W}$ is a mass parameter matrix determining the strength of the coupling and how the box can be deformed. The mass parameter matrix is given by

$$(\mathbf{W}^{-1})_{ij} = \frac{4\pi^2\beta_{ij}}{3\tau_p^2 L}, \tag{3.20}$$

where β_{ij} are the isothermal compressibilities, L is the largest box matrix element, and τ_p is the pressure time constant of the algorithm [153]. The Parrinello-Rahman algorithm is commonly used in pressure-coupling but may result in very large box oscillations if the system pressure is far from equilibrium. The Berendsen algorithm converges quickly and is therefore more suitable for the equilibration of the system,

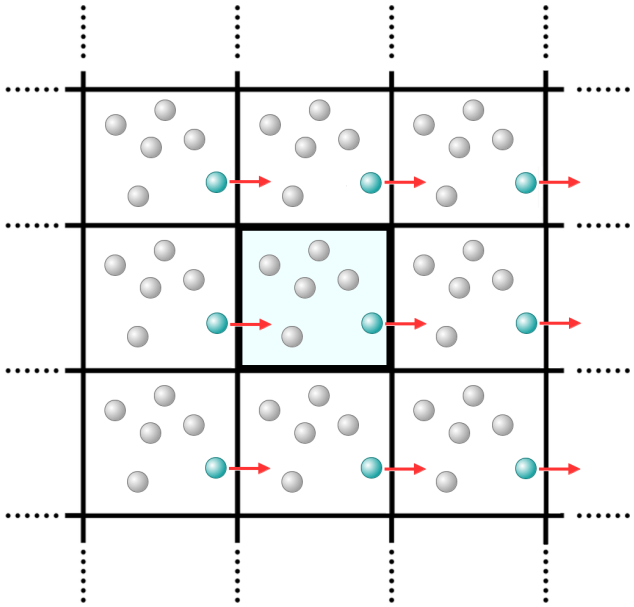


Figure 3.2: A schematic illustration of periodic boundary conditions.

after which the barostat is often switched to a more sophisticated one, such as the Parrinello-Rahman barostat [153].

3.4 Periodic Boundary Conditions Reduce Boundary Artifacts

An important property to be considered as well, is the overall shape of the system. Long-range interactions, especially those of electrostatic origin, highly depend on the boundary conditions applied to the system. The simulated system has some restrictions in size since it can not be infinite. Therefore the system boundaries should be carefully defined in order to avoid artifacts. A very simple way to overcome this problem is to apply periodic boundary conditions (PBCs) to the system.

The PBCs exactly replicate the system in three dimensions and provide a periodic lattice consisting of unit cells. When a particle moves out of the periodic simulation box, it instantly enters the box from the opposite side as illustrated in Fig. 3.2. Thus there are no boundaries of the system, only a space-filling box surrounded by translated copies of itself, thus the edge effect are minimized [159]. On the other hand, in a too small system the particles might interact with other particles more than once, which will cause undesired edge-effects. Therefore the size of the box must be chosen with care. The minimum image convention ensures that a given particle only interacts with the closest copy of another particle. Only the interaction with either a particle in questionper or its periodic image is taken into consideration, which one ever is closer.

3.5 Proceeding to the Simulation

A simulation study begins with the system preparation. For this purpose the initial conformations of the molecules are needed, provided for example by an X-ray structure or a theoretical model. Once all the needed structures are available, the system can be built by defining the positions for all atoms in each molecule. In addition to coordinates defining the positions, the structure of a system is defined by atomic velocities and by the forces acting on each atom. Atomic properties, such as their masses, charges, and bonded interactions must be defined in the topology of the system [153]. The size and the shape of the periodic box are also required before starting the simulation.

If the generated starting configuration of the system is far from equilibrium, some interactions might be high in energy and the MD simulation may fail due to this instability [153]. Therefore energy minimization is needed before starting the real simulation. The potential energy function of the system is very complex and has a global minimum as well as several local minima. The energy minimization finds the nearest local minimum of the potential function. The *steepest descents* method, that is possible to use in GROMACS, will bring the system potential function close to the nearest local minimum quite quickly by an algorithm that takes a derivative of the potential function in the direction of the negative gradient until it finds the local minimum [153]. In addition to the energy minimization, the system often needs some time for equilibration for example to stabilize the structure of a molecule, and therefore a comparatively short equilibration phase might be needed. After having the system energy minimized and equilibrated, one can proceed to the production phase of the simulation.

In general, MD simulations are based on an algorithm simplified in Fig. 3.3. The initial conditions, such as the positions and velocities, are first given to each particle. Next, the forces acting on the particles are calculated from the potential equation. After the forces, velocities and positions are known, the following step is to calculate new positions. The velocities are then calculated in these new positions with the leap-frog algorithm and scaled with thermostat and barostat to force the system into the correct ensemble. The time will be increased at this point, and if required, some information about the system such as the positions, velocities, potentials, temperature, etc. can be written down in the output file. This cycle is then repeated as many times as required until the desired amount of time steps has been reached. After finishing the simulation, the output files are used to analyze the results.

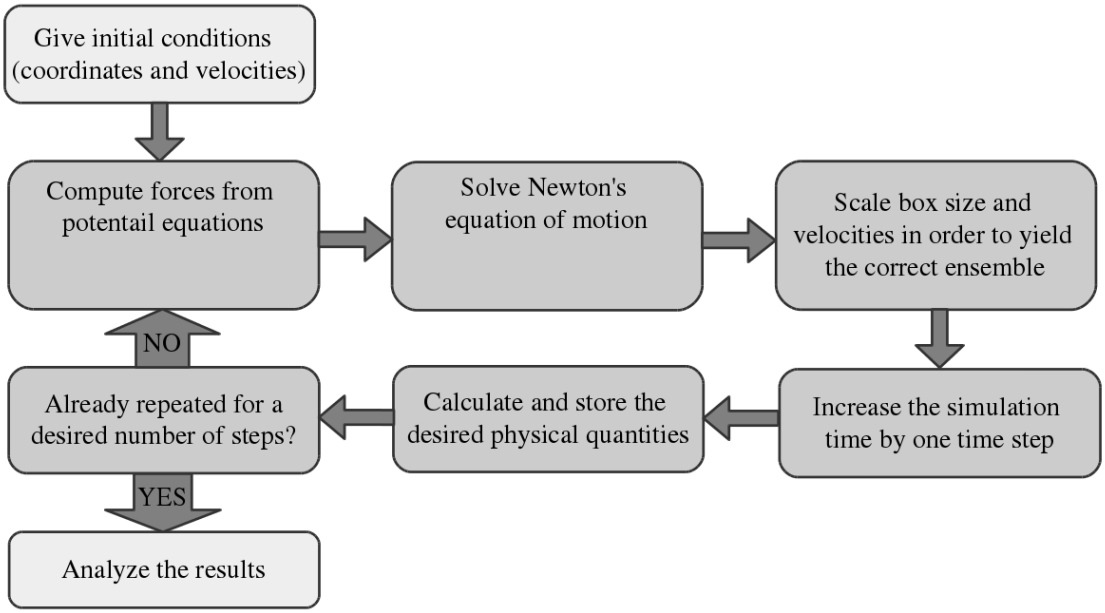


Figure 3.3: The simplified MD algorithm.

3.6 Limitations of Molecular Dynamics Simulations

The MD simulation method is a good approximation of real life events, yet it should be highlighted that it is based on a variety of approximations and assumptions that cause limitations for its use. Since the MD method is based on classical physics, the simulations work for most of the atoms at normal temperatures, but some processes, such as the proton or electron tunneling or chemical bonding, can not be properly treated by classical dynamics. MD simulations are restricted by the Born-Oppenheimer approximation which does not take the motions of electrons into consideration [153]. Electrons are supposed to adjust their atomic positions instantly upon a change in position, and remain in their ground state, for which reason the electron transfer processes and electronically excited states can not be treated [153].

Major limits for MD simulations are set by the simulation time scale and the system size due to computational power available. In the beginning of the 2010s, reasonable limit was a million particles corresponding to roughly about 5000 all-atom lipids, and simulation times of microseconds [151]. On the other hand, by going down in resolution from all-atom models to simplified coarse-grained models, the respective reachable limits were about 50 000 coarse-grained lipids, and time scales up to millisecond scale [151]. During the past few years, advances in software, hardware, and simulation techniques have even made milliseconds of atomistic simulations possible [162], and the accessible limits continue increasing.

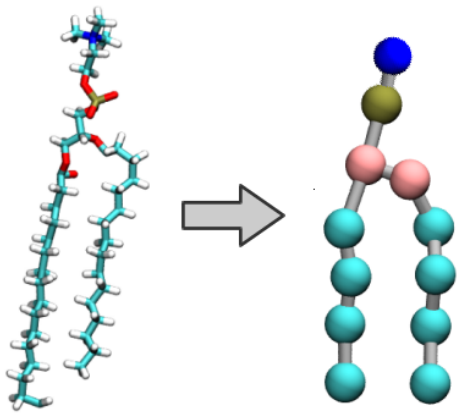


Figure 3.4: Representation of a lipid from all-atom to coarse-grained model.

3.7 Martini Coarse-Grained Force Field Provides a Simplification

The time scales of a variety of biologically interesting phenomena, such as dynamics of large proteins and the lipid phase separation, are currently challenging to reach with atomistic simulations. The Martini method developed for coarse-grained molecular modeling provides an alternative by allowing simulation times on length and time scales 2-3 orders of magnitude larger compared to atomistic simulations with the same computational effort [158]. This is done by simplifying the atomistic molecular models (Fig. 3.4). While in atomically detailed models the particles correspond to individual atoms in the system, in coarse-grained models several atoms are grouped into beads which reduces the number of force centers and thus the computational effort of the force evaluation [163]. The reduction of the number of force centers however leads to less friction between simulated molecules and thus speeds up their diffusion compared to atomistic systems [163]. Generally, the time in Martini simulations equals to around 4 times longer time in atomistic simulations [163, 164].

The Martini force field parameters have been developed for several biomolecules, including lipids, proteins, and DNA [158, 163, 165]. The force field parameters are not based on accurate knowledge of atomistic details, but instead the parameters can be used in a more general manner in a variety of different applications without the need to reparameterize by using the beads as building blocks.

The proteins in the Martini force field can be simulated in combination with an elastic network in order to help preserve their higher-order structure. An elastic network can be used as a structural scaffold to describe and maintain the overall

shape of a protein by adding extra harmonic bonds to the standard Martini topology between non-bonded beads based on a distance cut-off. The resulting structural and dynamical properties of a protein in MD simulations with an applied elastic network are comparable to those obtained using atomistic protein models [166]. Notable is that conformational changes necessary for protein folding can not be expected to produce because of the elastic network's built-in structural bias toward the reference configuration [166].

Since the chemical resolution of a coarse-grained system is limited, some properties of varying importance are lost. For some phenomena, small changes such as tilting of an individual lipid, are not important in the overall picture, such as phase separation of lipids, yet this might be critical in, for example, protein-lipid interactions. Therefore the conversion of a coarse-grained system on atomistic model is sometimes necessary to obtain insight about atomistic–scale processes from coarse-grained simulations. The conversion from a coarse–grained to an atomistic model and vice versa is possible with several scripts available [165, 167, 168].

4. SIMULATION MODELS AND ANALYSIS METHODS

The aim of this study was to find out the preference of β_2AR for domains in a phase segregated bilayer. From the computational perspective, it is extremely important, first of all, to choose correct models for the molecules in order to correctly mimic the underlying biology. Next, the system settings from the point of view of the study must be chosen, while aiming to get results reliable enough and balancing with the computational power available. Finally, the obtained simulations must be analyzed with care, and the most informative ways to present the significant results from the massive amount of data must be discovered. In this chapter the preparation procedure of the studied systems is presented, along with the simulation parameters and details, and the analysis methods used.

4.1 Systems

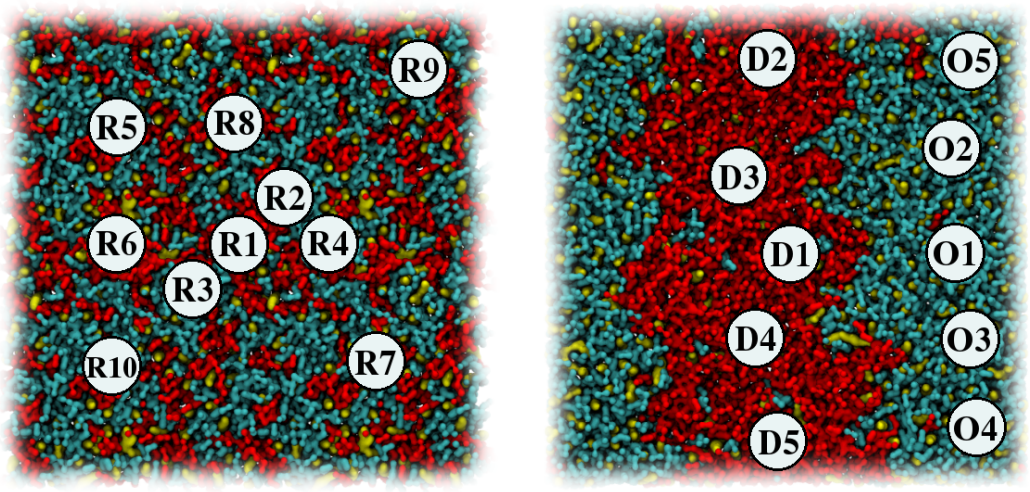
The lipid bilayer simulated by Risselada and Marrink [145] was used as a platform for the simulations. The lipids dipalmitoyl-phosphatidylcholine (diC(16:0)PC, DPPC), dilinoleoyl-phosphatidylcholine (diC(18:2)PC, DLiPC), and cholesterol (CHOL) were used with the DPPC:DLiPC:CHOL ratio being 42:28:30, and the total number of lipids of 2200. This lipid mixture was found to spontaneously form a strongly phase separated bilayer composed of a disordered L_d domain and an ordered L_o domain [20]. The CHARMM-GUI graphical user interface [169] was used for building a patch of the lipid bilayer with the bilayer builder in the Martini maker options [170]. The bilayer contained 231 DPPC, 154 DLiPC, and 165 CHOL molecules, consisting of 550 lipids in total. 50 Å of water was added on both sides of the bilayer, totaling $\sim$44000 coarse–grained water beads, along with 1040 Na and Cl ion molecules to reach the physiological salt concentration of 0.15 M. The system was equilibrated and multiplied by 4 in order to obtain the bilayer of 2200 lipids. The protein was then embedded in ten different locations yielding ten systems, R1-R10, with different starting structures. The GROMACS tool `gmx solvate` was used to insert the protein in the bilayer and to remove the overlapping lipid, water, and ion molecules.

After embedding the protein, the total number of lipids was adjusted to 2000, the same as in Risselada's and Marrink's study [145], and the number of ions was modified in order to balance the charges in the system that now contained protein. The locations of the protein in each system are shown in Fig. 4.1(a). Another two systems, RC1 and RC2, were constructed by embedding the protein in the bilayer in the same locations as in systems R1 and R9, respectively. In these systems the CHOL parameter developed by M. Daily et al., discussed in more detail in the next section, was used [171].

The lipid bilayer was simulated for 5 μs until the phase separation of lipids had occurred. This system was then used as a platform to embed the protein with a similar procedure in ten different locations, yielding five systems, O1-O5, with the protein initially in the L_o phase, and five systems, D1-D5, with the protein initially in the L_d phase. The locations of the protein embedded to the phase separated bilayer are shown in the Fig. 4.1(b), and a side view of the system in Fig. 4.2.

The atomistic systems were constructed with backmapping script provided in reference [167]. Approximately 1/4 of the lipids around the protein in the systems R1 and D1 were backmapped from coarse–grained to atomistic systems B1 and B2, respectively. The simulation of an atomistic system is computationally more costly, and therefore only 1/4 of the coarse–grained systems was backmapped to atomistic detail. The system R1 was backmapped after the 6 μs of simulation, yielding the system B1 with the lipid composition listed in Table 4.1. In this system the protein was in the boundary of the L_o and L_d phases. The initial structure of the system D1 after equilibration, where the protein was entirely in the L_d phase, was backmapped yielding the system B2 with lipid composition listed in Table 4.1 as well. These configurations were chosen for backmapping in order to enable the protein to spontaneously create its preferred environment in the atomistic simulations. The original lipid ratio was not the same as in Martini systems since the protein location dictated the lipids to be backmapped. The DLiPC molecules were backmapped into dioleoyl-phosphatidylcholine (diC(18:1)PC, DOPC) due to the force field parameters available. DOPC is an unsaturated lipid like DLiPC, but has only one double bond in each chain instead of two.

The backmapped systems were equilibrated first with position restraints on all protein atoms, and on lipid head groups and chains in the z-direction (direction perpendicular to the bilayer plane) until the bilayer filled the simulation box in the x- and y-dimensions. In both systems only the lipids were backmapped, and water and ions were added after the equilibration of the box. The systems were



(a) Protein embedded in S1 (b) Protein embedded in S2

Figure 4.1: A top view of the simulation box of (a) the bilayer with random lipid distribution, in which the protein was embedded to generate different starting structures for ten individual systems R1-R10, and (b) the phase separated bilayer, in which the protein was embedded to generate five different starting structures with the protein in the L_o phase (O1-O5) and five with the protein in the L_d phase (D1-D5). The DPPCs are shown in cyan, forming the L_o phase together with CHOL that is shown in yellow, and the L_d phase is mainly composed of DLiPCs shown in red. The circles indicate the locations in which the protein was embedded in the system in question.

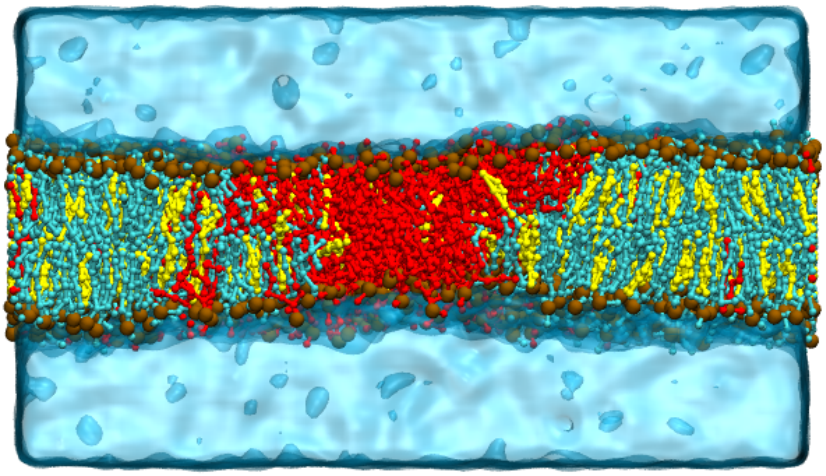


Figure 4.2: A side-view of the simulation box of the phase-separated bilayer. Water is shown in blue, facing the bilayer from upper and lower sides. CHOL molecules are shown in yellow, DPPC in cyan, DLiPC in red, and the phosphorus containing groups are shown in ochre.

Table 4.1: The system compositions. The protein was placed in the lipid bilayer containing 2200 lipids, which yielded systems R1-R10, O1-O5, and D1-D5 that differ in protein location. In systems RC1 and RC2 the protein was embedded at the same locations as in R1 and R9 respectively, but was simulated with a different CHOL parameter. Systems B1 and B2 are smaller, atomistic systems, backmapped from systems R1 and D1, respectively.

Molecule	**Number**	**R1-R10,O1-O5, D1-D5, RC1,RC2 (coarse-grained model)**	**B1** (all-atom model)	**B2**
β_2AR	-	1	1	1
DPPC	924	840	193	183
DLiPC	616	560	-	-
DOPC	-	-	153	169
CHOL	660	600	148	143
Lipids in total	2200	2000	494	495

Table 4.2: The protein initial location in the constructed systems, CHOL parameters used, and the simulation times of the systems.

Forcefield	**System**	**Protein location**	**CHOL parameter**	**Run time (μs)**
Martini	Lipid bilayer	-	[172]	5
	R1-R10	Random	[172]	10
	O1-O5	L_o phase	[172]	10
	D1-D5	L_d phase	[172]	10
	RC1, RC2	Random	[171]	10
OPLS-AA	B1	L_o/L_d border	[173–175]	2
	B2	L_d phase	[173–175]	2

then equilibrated without any restraints on the lipids, but keeping the ones on the protein atoms until the bilayer had stabilized. The protein restraints were then released in steps by first equilibrating the system for 10 ns with position restraints on the protein heavy atoms, and then for 20 ns with position restraints on the protein backbone. After the equilibration, the systems were run for 2 μs, as listed in Table 4.2.

4.2 Molecular Models of the Structures

The initial structure of the human β_2AR (residues 32-342) in the OPLS-AA force field used in the simulations was obtained from the PDB id 3D4S [18], with modifications discussed in [174]. The fine grained structure of the protein in OPLS-AA

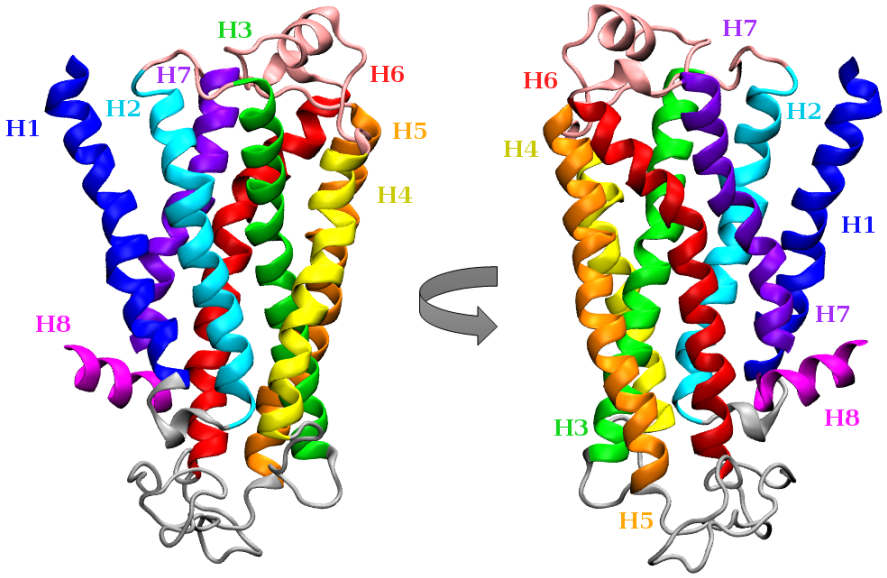
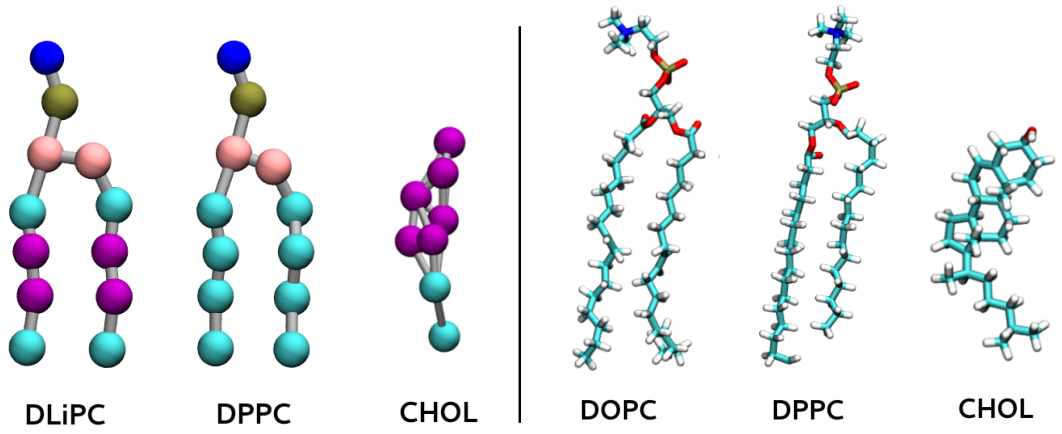


Figure 4.3: The structure of β_2AR used in this study [174]. The transmembrane helices are colored for clarity: helix 1 (blue), helix 2 (cyan), helix 3 (green), helix 4 (yellow), helix 5 (orange), helix 6 (red), helix 7 (purple), and a short helix 8 (magenta). Intracellular loops are shown in grey and extracellular loops in pink.

force field was converted into a coarse-grained one, to the Martini 2.2 force field, by using the script *martinize.py* provided by the Martini development team [165, 168]. An elastic network was used for the coarse-grained protein with an elastic bond force constant of 500 kJ·mol^{-1}·nm^{-2}, with lower and upper elastic bond cut-offs of 0.5 and 0.9 nm, and with bond strengths independent of the bond length. The structure of β_2AR is shown in Fig. 4.3.

The size of the Martini coarse-grained cholesterol size has been found to have a significant impact on the behavior of cholesterol-containing bilayers [176]. A cholesterol parameter with a virtual site description developed by M. Melo, H. Ingólfsson, and S. Marrink [172] was used in systems R1-R10, O1-O5, and D1-D5 instead of the earlier Martini cholesterol model [158] that has been reported to drive the L_o phase too ordered and fail to preserve its fluidity [177, 178]. The cholesterol parameter of Melo & al. was developed by reparameterizing the earlier Martini cholesterol model, by adjusting its packing properties to a allow a proper L_o cholesterol behavior, and stabilizing it by reformulating the spatially constrained cholesterol topology using virtual interaction sites [172]. Another cholesterol parameter for Martini developed by M. Daily et al. was reported to better mimic the properties of atomistic choles-



(a) Coarse–grained lipids (b) Atomistic lipids

Figure 4.4: (a) The coarse-grained models of the lipids, and (b) the atomistic models of the lipids. The beads of coarse-grained PC lipids are colored separately: nitrogen group (N) in blue, phosphate group (PO4) in pickle green, glycerol groups (GL) in pink, and chain carbon groups in purple if containing double bonds (D), otherwise in cyan (C). The beads for cholesterol ring structure are shown in purple and the hydrocarbon chain beads in cyan. The atoms of the fine–grained lipids have similar color codes: nitrogen is shown in blue, phosphorus atoms in pickle green, oxygen in red, carbon in cyan, and hydrogen in white.

terol [171], and was used in systems RC1 and RC2. The Martini coarse-grained force field parameters were used in all Martini systems for saturated DPPC lipids and for the unsaturated DLiPC lipids, as well as for water and ions [158, 163, 179]. Fig. 4.4 shows the coarse–grained representations of the DPPC, DLiPC, and CHOL molecules. The lipid parameters in the all-atom OPLS force field were taken from references [180–182] for DPPC and DOPC, from references [173–175] for CHOL, and from [174] for β_2AR. The TIP3P water model, compatible with the OPLS parameterization, was used for water [183].

4.3 Simulation Details

All simulations were carried out with the GROMACS 5.0.x MD package [184]. The coarse-grained systems (R1-R10, RC1-RC2, O1-O5, D1-D5) were simulated with the Martini 2.2 force field [163, 165] with a 20 fs integration step. In all simulations the solvent molecules, lipids, and the protein were independently coupled to a constant temperature bath [185] with a time constant of $\tau_T = 1.5$ ps at a temperature of 295 K that was also used in Risselada's and Marrink's simulations [145]. At this temperature the liquid domains were observable and stable. The pressure was weakly

coupled [185] to 1 bar with a time constant of 3.0 ps. The semi-isotropic coupling of the pressure was applied in the xy-plane (the membrane plane), and independently in the z-direction. The Coulomb interactions were treated with a Coulomb potential that is screened with a relative dielectric constant of $\epsilon_r = 15$. The Lennard-Jones potential was smoothly shifted to zero between 0.9 nm and 1.1 nm and the Coulomb potential between 0 nm and 1.1 nm.

The lipid bilayer with a random lipid distribution obtained with CHARMM-GUI was energy minimized for 5000 steps and then equilibrated for 10 ns with a time step of 20 fs. After the protein was embedded in each coarse-grained system, the systems were energy minimized for 5000 steps and then equilibrated for 4 ns, first with position restraints on all protein beads, lipid head groups and chains, and on water and ion molecules in z-direction, and finally equilibrated without any restraints for 10 ns. The systems were then run for 10 μs each. The simulation times are listed in Table 4.2. The simulation time of the coarse-grained systems should be scaled by using a standard conversion factor of 4 [163]. The total effective time sampled is therefore $\sim$40 μs for each coarse-grained system.

The OPLS-AA force field [157] was used in the systems B1 and B2 with a 2 fs time step. The simulations were performed at 295 K as well by employing the v-rescale thermostat [186] with a time constant of 0.1 ps. The pressure was coupled to 1 bar with the Parrinello-Rahman barostat [187] and with a time constant of 1 ps. The Lennard-Jones interactions were cut off at 1.1 nm. For the long-range electrostatic interactions, the particle mesh Ewald method (PME) [188] was employed with a real space cut-off of 1.0 nm, β-spline interpolation (order of 6), and a direct sum tolerance of 10^{-6}.

The simulations were performed in the isobaric-isothermal (NpT) ensemble. In each of the systems, PBCs were applied in three dimensions. All the systems were energy minimized using the steepest descent algorithm. The LINCS algorithm was used to constrain the bonds between the atoms [189]. The neighbor search algorithm, the grid, updated neighbor lists every 10 time steps. Visual images of the systems or their compounds were prepared using the VMD (Visual Molecular Dynamics) software [190] version 1.9.1. The simulations were performed in the Center for Scientific Computing (CSC) in Finland and in the Tampere Center for Scientific Computing (TCSC).

4.4 Analysis Methods

Many tools are available for the simulation data analysis, for instance some built-in tools provided by the GROMACS software package [153] or other tools freely available. However, some tools must be developed further or from scratch in order to present the data in a specific form. The most informative analysis tools used in this study are presented below. All analyses were performed with a time resolution of 200 ps and for the last 5 μs trajectory of each coarse-grained system unless mentioned otherwise. For atomistic systems, a time resolution of 100 ps was used and the last 1 μs trajectories were chosen for the analyses.

4.4.1 Fraction of Lipid–Lipid and Protein–Lipid Contacts

A contact is considered to take place between two atoms or molecules if the distance between them is less than a certain cut-off distance. The fraction of lipid–lipid contacts was calculated by dividing the number of contacts between the saturated (DPPC) and unsaturated (DLiPC or DOPC) lipids by the total number of contacts between unsaturated and all PC lipids (DPPC and DLiPC/DOPC). The fraction of the β_2AR–lipid contacts was calculated by dividing the number of β_2AR–unsaturated lipid contacts by the number of β_2AR–all PC lipid contacts. The number of contacts was calculated with GROMACS tool `gmx mindist` [153], and the cut-off distance used was 0.6 nm.

4.4.2 Concentration of Lipids

The local concentration of any lipid type in the membrane was calculated by first taking the last 1 μs trajectory of each system and then calculating the center of mass (COM) coordinates of each lipid with the GROMACS tool `gmx traj` [153]. The membrane was then divided into 20 slabs, of approximately 1.1 nm of width, along the axis perpendicular to the phase boundaries. The number of each lipid type in each slab was then calculated based on the COM coordinates of the lipids. The number of the lipids in question was divided by the total number of lipids in each slab in order to obtain the fraction of the lipids in that slab.

The concentration of lipids in the membrane at certain distances from the protein was calculated similarly. The lipids where divided into 20 groups based on the distance between the COM of each lipid and protein. The distances defining the groups were $[0., 1.5)$, $[1.5, 2.0)$, $[2.0, 2.5)$, $[2.5, 3.0)$,...,$[9.5, 10.0)$. The fraction of each lipid type in a certain group was then calculated by dividing the number of lipids

in question by the number of lipids in the group.

4.4.3 Radial Distribution Function

The radial distribution function (RDF) is used to determine the distribution of particles around a reference particle. RDF or in other words the radial distribution $g_{AB}(r)$ between particles of type A and B is defined by [153]

$$\begin{aligned} g_{AB}(r) &= \frac{\langle \rho_B(r) \rangle}{\langle \rho_B \rangle_{local}} \\ &= \frac{1}{\langle \rho_B \rangle_{local}} \frac{1}{N_A} \sum_{i \in A}^{N_A} \sum_{i \in B}^{N_B} \frac{\delta(r_{ij} - r)}{4\pi r^2} \end{aligned} \tag{4.1}$$

where N_A and N_B are the number of particles of type A and B, respectively, $\langle \rho_B(r) \rangle$ is the particle density of type B at a distance r around the particle A, and $\langle \rho_B \rangle_{local}$ the particle density of type B averaged over all spheres around the particle A. The calculation of RDF is based on an algorithm that determines the number of particles within a distance between r and $r + \mathrm{d}r$ from a reference particle for every r. The results can be plotted giving the density of certain particles as a function of the distance. The RDF can be regarded as a measure of the probability of finding a particle at a distance r from the reference particle.

In this study the analysis tool `gmx rdf` provided by GROMACS [153] was used to calculate the radial distribution function of the COM of the lipid molecules from the protein in x- and y-dimensions. PBCs were used to calculate the distances, and the trajectories were divided into 2 μs intervals for which the RDF was calculated in order to find out the time evolution.

4.4.4 Two-Dimensional Number Density

A two-dimensional (2D) number density map shows the spatial distribution of lipids in the bilayer plane. For this analysis, the protein molecules were first centered in the simulation trajectory, and the rotational and translational motion of the protein were removed from the analysis. With the GROMACS tool `gmx traj` [153] the center of mass coordinates of each lipid were calculated separately, as well as the coordinates of the upper and lower parts of the protein TM helices. The trajectories of the systems with a similar starting configuration were merged into longer trajectories, the systems with an initially random lipid distribution (R1-R10) were concatenated, and this was also done for systems with the protein initially in the L_o phase (O1-O5),

for the systems with the protein initially in the L_d phase (D1-D5), and for systems RC1-RC2.

The box dimensions were calculated for each trajectory with the GROMACS tool `gmx energy`, and the greatest box dimension was used to divide the box into a grid in such a way that one spatial step in the grid corresponds to approximately 1 Å. The average of the COM coordinates of the lipids in the z-direction were used to define the center of the bilayer at each time step. The lipids of each type (DPPC, DLiPC/DOPC, CHOL) were then divided into the upper and lower leaflet based on their z-coordinate value, yielding six density maps for each trajectory. The contribution of each lipid was added to the corresponding grid point in the density map. The coordinates of the COM of each lipid were also checked to be inside the box in order to get the contribution of each lipid to fit in the density map. Finally, the maps were normalized by dividing the number of frames in the trajectory and by the number of lipids in question. Similar density maps were also calculated for the upper and lower parts of the protein TM helices, and the location of the protein helices was added to the obtained lipid number density maps.

4.4.5 Lipid Occupancy Time per Residue

The number of contacts between each lipid type (DPPC, DLiPC/DOPC, CHOL) and each protein residue were averaged over the systems R1-R10, RC1-RC2, O1-O5, and D1-D5. The residues were considered to be in contact with a lipid if the distance between them was equal to or less than 0.6 nm. The `-group` option of the GROMACS tool `gmx mindist` was used to count the contacts of multiple lipids of the same type with a protein residue as one contact at one time step. The total occupancy time was normalized in such a way that the occupancy time of 1 indicated a contact that lasted throughout the simulation, whereas the value of 0 means lack of contacts. The total occupancy time per protein residue was plotted on the protein surface separately for each lipid type with VMD.

4.4.6 Lipid Tail Order Parameter and Bilayer Thickness

The order of the lipid bilayers is characterized by an order parameter that can be measured by deuterium NMR. The order parameter describes the average orientation of the carbon chains and describes how straight or twisted they are. For lipid chains in atomistic systems the order parameter is described by the equation [153]

$$S_z = \frac{3}{2}\langle \cos^2 \theta_z \rangle - \frac{1}{2}, \tag{4.2}$$

where θ_z is the angle between the z-axis of the simulation box and the vector between two sequential carbon atoms in the lipid chain. The squared cosine values of the angles are averaged over time and molecules in order to obtain the order parameter S_z that has a value between -0.5 and 1, where the value of 1 implies full order perpendicular to the normal. The order parameters calculated for the Martini coarse-grained lipids with this method do not compare to any values obtained in atomic systems, since multiple chain carbons are considered as one bead, and the order parameter values are obtained between the beads that are segments in the chain. However, the order parameter values between different Martini coarse-grained lipids are comparable, so it gives some hint of the differences in the order between for example regions in one system or two completely different systems.

In this study the order parameters of the Martini coarse-grained lipid chains were calculated using Eq.(4.4.6), with the script `do-order-multi.py` provided by the Martini development team [163]. The chain beads GL1-C4A and GL2-C4B of the Martini coarse-grained PC lipids were chosen for the analysis. The local lipid chain order parameter and bilayer thickness were calculated in the xy-plane using the tool `g_lomepro` [191], which analyzes the local properties of the membrane (lipid tail order parameters, membrane thickness, area per lipid, membrane curvature) in two or three dimensions based on a grid. The bilayer thickness was calculated as the average distance in the z-direction between the phosphorus containing groups of the PC lipids in the upper and lower leaflets. The systems with a similar starting structure were concatenated into same trajectory for the calculation.

4.4.7 Diffusion Coefficient

The diffusion coefficient D_A of the molecules of type A in the bilayer can be determined by first calculating the mean squared displacement (MSD) $\langle ||r_i(t) - r_i(0)||^2 \rangle$ of all molecules i in the group A. To determine the coefficient D_A, one can use the

Einstein relation [153]

$$\lim_{t\to\infty} \langle ||r_i(t) - r_i(0)||^2 \rangle_{i\in A} = 4D_A t, \tag{4.3}$$

which says that in a two-dimensional system, a four-fold diffusion coefficient D_A is linearly comparable to the long–time limit of the slope of the MSD. When calculating the diffusion in three dimensions, instead of two as in this study, the integer before D_A becomes 6.

The GROMACS tool `gmx msd` [153] was used to calculate the MSD of the lipids, and the diffusion coefficient was determined from the slope of the MSD plotted as a function of time. Only the linear part of the curve was taken into account for the calculation of slope. For Martini coarse-grained systems the diffusion has been reported to be too fast compared to the atomistic systems [163], and thus the obtained diffusion constant has to be divided by an additional factor of 4. The jumps of the molecules in the periodic boundaries and the COM motion were removed from the trajectories for the analysis.

5. RESULTS AND DISCUSSION

In this chapter the findings regarding the partitioning of β_2AR in a highly phase separable lipid bilayer are presented. A total of 22 coarse-grained MD simulations were carried out with different starting structures and cholesterol parameters, as well as two fine grained systems in order to justify the results obtained from coarse-grained simulations. The systems and their compositions are listed in Table 4.1 and 4.2. The presented graphs have been plotted with MATLAB or directly by the LaTeX platform with PGF/TikZ. Snapshots of the simulated systems have been prepared with VMD.

5.1 β_2AR Partitions to the L_o/L_d Phase Boundary

The phase separation can be followed in time by calculating the fraction of lipid–lipid contacts, shown in Fig. 5.1(A) for systems R1-R10 with an initially random lipid distribution around the receptor. Within 4-5 μs the fraction of contacts stabilizes in each system, indicating that the lipids have found their equilibrium location and separated into two phases. In systems O1-O5 and D1-D5 the protein was placed in an already phase separated bilayer. The protein finds its position with respect to lipids within approximately 3 μs as the fraction of β_2AR–DLiPC contacts shows in Fig. 5.1(B). For each system, the fraction of protein–unsaturated lipid contacts stabilizes after 3 μs and the fraction stays at a value of $\sim$0.8. In case the protein is initially placed in the L_o phase (Fig. 5.1(C)), it takes a longer period of time, up to 5 μs for stabilization, until the fraction of contacts reaches the value of 0.8. The fraction of protein–lipid contacts stabilizes within a few μs in systems in which the protein was initially placed in the L_d phase (Fig. 5.1(D)), also exhibiting a contact fraction of 0.8 between the protein and unsaturated lipids. The systems can thus be considered to be stable enough regarding the phase separation and protein location, when simulated for 10 μs. Notable is that after equilibration the unsaturated lipids are dominating the protein-lipid contacts with almost 80 % of lipid contacts, irrespective of the starting configuration.

During the 10 μs simulation of the coarse-grained systems with different starting structures, β_2AR partitions to the phase boundary in each case. Snapshots from the

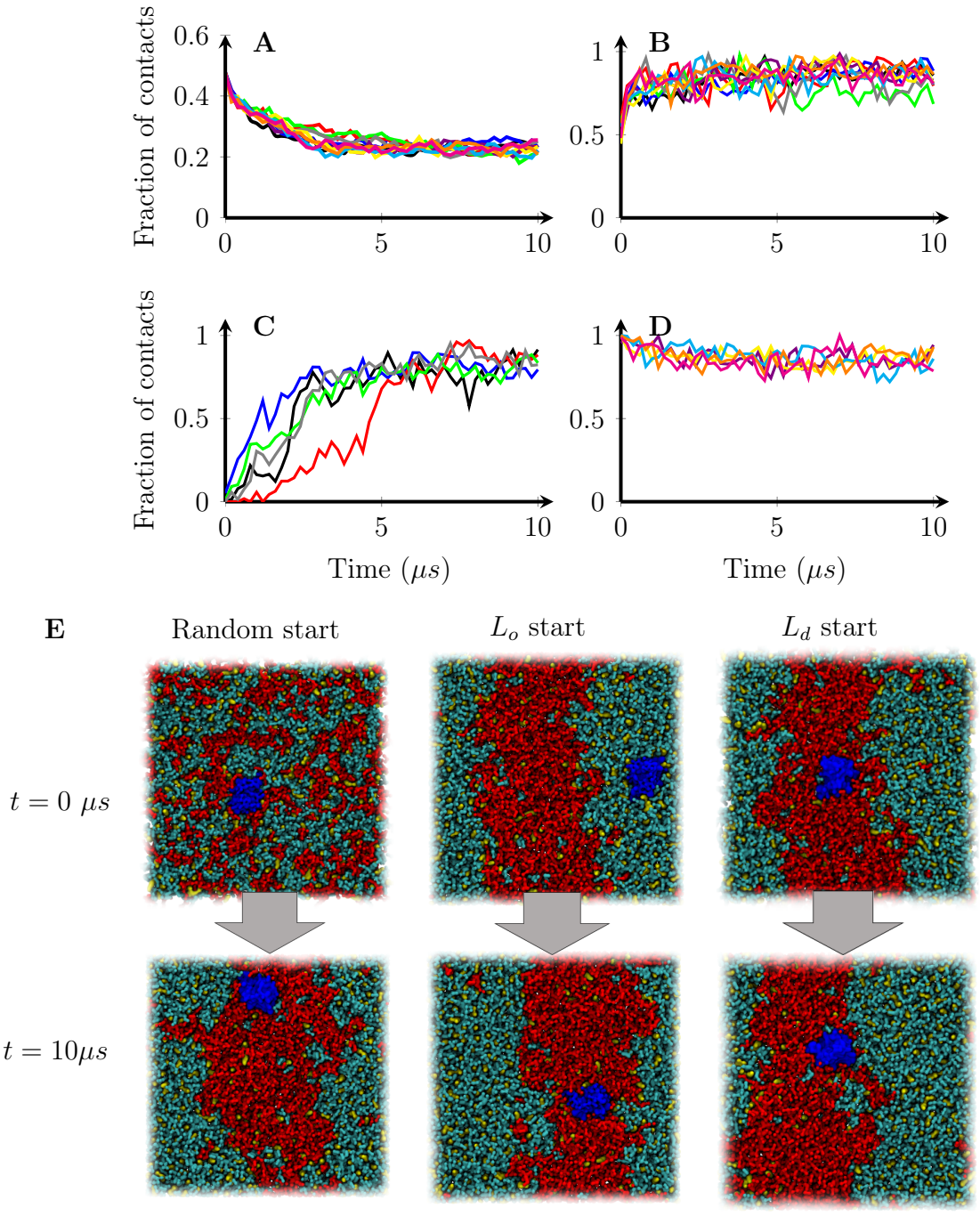


Figure 5.1: (A) The fraction of DLiPC-DPPC contacts in systems R1-R10, and the fraction of β_2AR-DLiPC contacts in systems (B) R1-R10, (C) O1-O5, and (D) D1-D5. Different colors correspond to different systems. E) Snapshots from the beginning of the simulation (top) and after 10 μs (bottom) in systems R1 (left), O1 (center), and D1 (right). Color codes in (E) are the same as in Fig. 4.1 and Fig. 4.2.

beginning and after 10 μs of simulation of systems R1, O1, and D1 are shown in Fig. 5.1(E) as an example of the β_2AR behavior when placed initially in a random lipid distribution, in the L_o phase, and in the L_d phase. The snapshots after 10 μs of simulation in all systems establish a similar location for the protein. Even though β_2AR partitions to the phase boundary, in the snapshots after 10 μs it is always mostly surrounded by polyunsaturated DLiPC lipids. In some cases, as shown after 10 μs in system O1, the L_d phase gets even narrower around the protein, allowing it to face the phase boundary at both sides. Since the systems are considered stable after 10 μs of simulation and the protein in each case prefers a similar location, these results indicate that the protein has found its most preferred location in the studied membrane.

The lipid distribution between the two phases can be seen in a more quantitative manner by plotting the concentration of lipids along the axis perpendicular to the phase boundary. The lipid concentrations in the two phases are shown in Fig. 5.2(A) in system R1 during the last 1 μs of the simulation. The raft–like L_o domain is composed of 60 % of saturated DPPC, 38 % of cholesterol and 2 % of unsaturated DLiPC, forming almost a pure binary mixture of saturated lipids and cholesterol. The mole fraction of cholesterol is even higher than the overall cholesterol concentration (0.3) in the membrane, and yet below the solubility limit of 0.66 for cholesterol in a DPPC membrane [192]. The L_d domain is composed of 8 % of DPPC, 14 % of cholesterol, and 78 % of DLiPC. The cholesterol mole fraction is thus reduced to half of the total cholesterol concentration, showing that the L_d phase is comparatively cholesterol–poor. The obtained concentrations are similar to the ones obtained by Risselada and Marrink [145] and in agreement with the composition of phase separated membranes observed in NMR measurements [59]. The same figure shows the location of the protein COM plotted along the same axis. The protein preferably stays on the side of the L_d phase, yet not completely in the center of the phase, but instead close to the phase boundary. The concentration of lipids and the location of the protein COM along the axis perpendicular to the phase boundary were similar in each system.

The concentration of lipids at increasing distances from the protein averaged over different types of systems is plotted in Fig. 5.2(B-D). The concentrations show a small peak of CHOL at the distance of 1.5 to 2.5 nm. The concentrations closer than 1.5 nm from the β_2AR COM show more differences than the concentrations further from the protein since the amount of lipids within such a short distance from the protein COM is small. Nevertheless, the most predominant lipid is the unsaturated

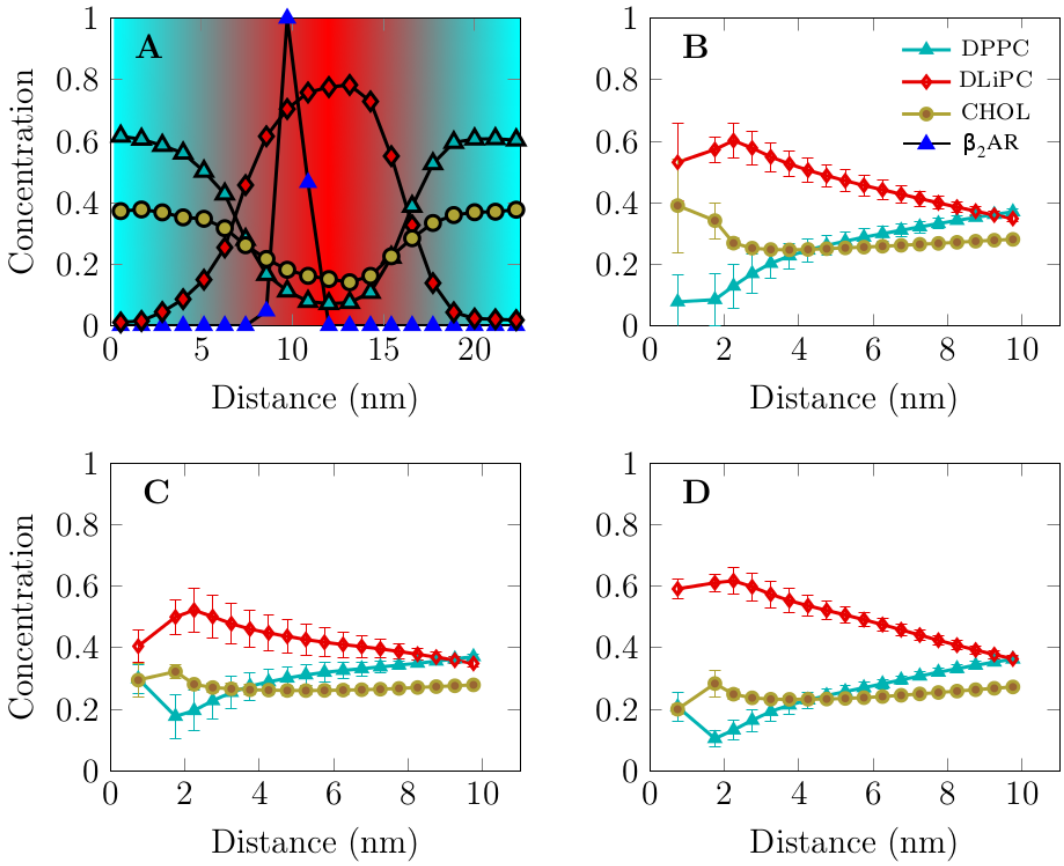


Figure 5.2: (A) The lipid concentration along the axis perpendicular to the lipid phase boundary in the system R1. The respective plots are similar in all systems, only the location of the phases changes. The distribution of the protein location is plotted as well along the same axis. (B,C,D) The concentration of lipids at increasing distances from β_2AR averaged over systems (B) R1-R10, (C) O1-O5, and (D) D1-D5.

DLiPC in each type of system. At distances of 1.5 to 2.5 nm, the DPPC fraction varies from 0.08 (R1-R10) to 0.18 (O1-O5), the fraction of DLiPC from 0.49 (O1-O5) to 0.59 (R1-R10 and D1-D5), and the fraction of CHOL from 0.29 (D1-D5) to 0.33 (O1-O5 and R1-R10). The cholesterol concentration close to the protein is similar to the one in the phase boundary in Fig. 5.2(A), yet the DLiPC concentration is higher and the DPPC concentration lower compared to the phase boundary concentrations. The values indicate that β_2AR preferentially resides in the phase boundary, slightly more on the side of the L_d phase, but has a higher concentration of CHOL around it. This is in line with the findings that have shown β_2AR to preferably interacts with CHOL [16–18].

5.2 Time Evolution of Lipid Radial Distribution Function from the β_2AR

The time evolution of the DLiPC, DPPC, and CHOL RDFs are plotted in Fig. 5.3 along with the final RDF of all lipids in systems R1, O1, and D1. In the case of an initially random lipid distribution of system R1, the RDF of DLiPC shows a clear peak that increases in time, whereas the amount of DPPC close to the protein decreases in time. The distribution of CHOL is also higher close to the protein COM during the first two microseconds, but decreases slightly and stabilizes after that, showing a peak within 2 nm from the protein. When the protein is placed in the L_o phase in system O1, there are basically no DLiPC molecules close to the protein, but the distribution near the protein increases rapidly in time and shows a similar peak as in the case of the system R1. The distributions of DPPC and CHOL are higher close to the protein COM in the beginning than after 10 μs of simulation, decreasing within 4 μs to similar curves as in the system R1. The RDFs of lipids do not vary much in time in the system D1 when the protein is placed in L_d, and within a few microseconds stabilize to similar distributions as in the final state of the systems R1 and O1. This implies that the protein is closest to its equilibrium location with respect to the phases when it is initially placed in the L_d phase.

The radial distributions of lipids from the protein stabilize to similar distributions within four microseconds in each system, showing similar behavior of the lipids surrounding the protein irrespective of the starting configuration. The final lipid distributions show that DLiPC molecules are the dominating ones within 2 nm from β_2AR, only a small amount of CHOL molecules reside closer to β_2AR. At the location of the peak in the DLiPC radial distribution, there is also a smaller peak in the CHOL distribution. The DPPC molecules are basically all located further than 2 nm from the protein, implying that β_2AR is preferably surrounded by unsaturated DLiPC lipids and cholesterol.

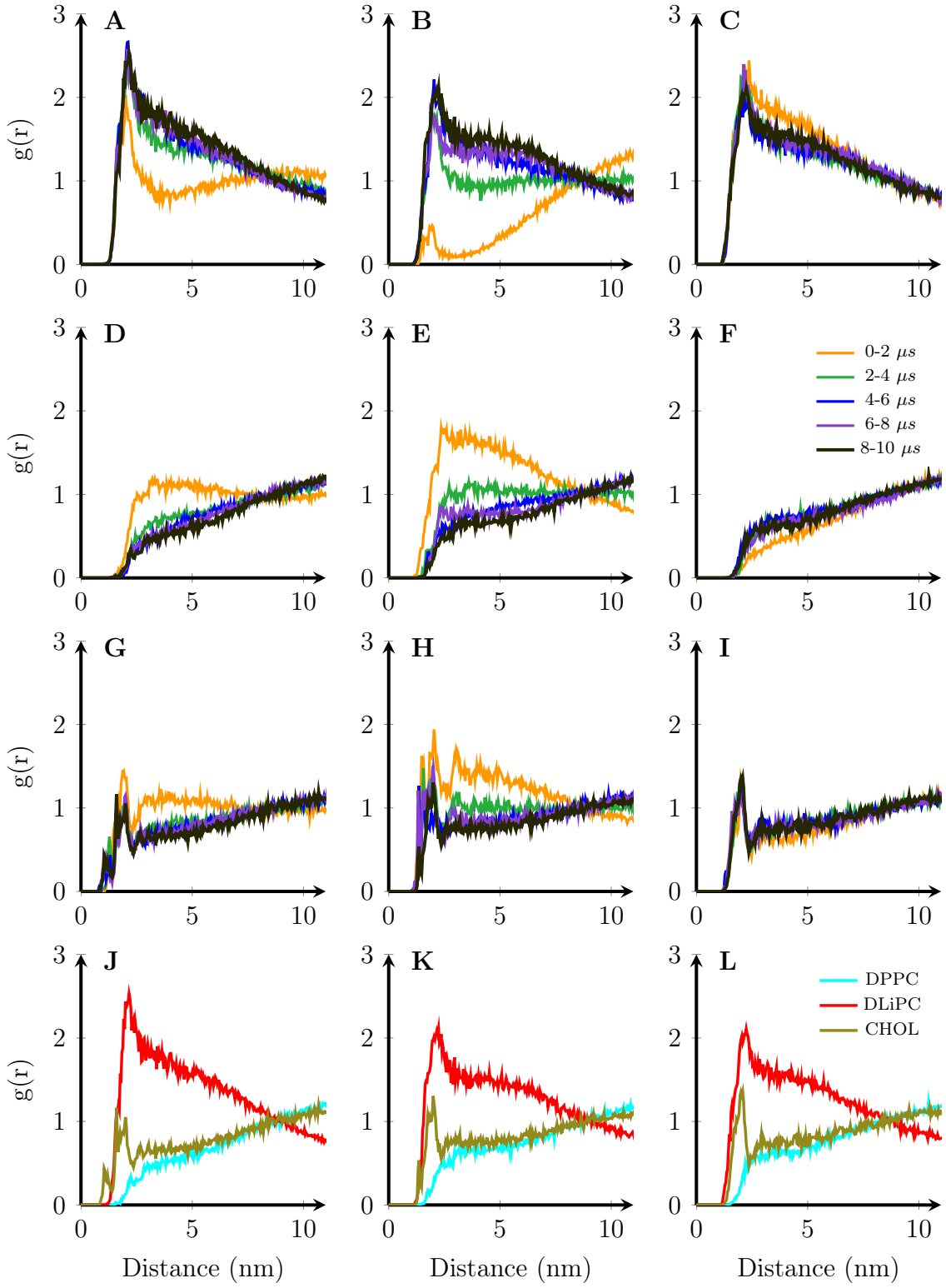


Figure 5.3: The time evolution of (A-C) DLiPC, (D-F) DPPC, and (G-I) CHOL RDF in systems R1 (left: A,D,G), O1 (center: B,E,H), and D1 (right: C,F,I). (J-K) The final RDF of all lipids in the respective systems.

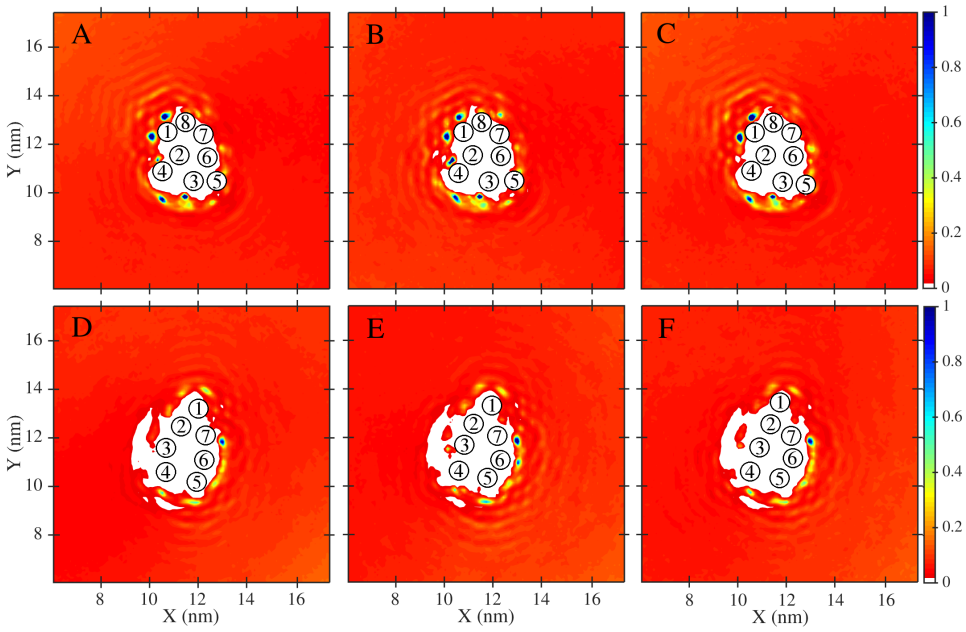


Figure 5.4: (A-C) The number density of CHOL around β_2AR on the intracellular leaflet of the bilayer and (D-E) on the extracellular leaflet of the bilayer, averaged over systems (A,D) R1-R10 (B,E), O1-O5, and (C,F) D1-D5.

5.3 β_2AR Is Surrounded by Unsaturated Lipids and Cholesterol

The spatial number density maps of the lipids in the xy-plane averaged over systems with similar starting configurations are shown in Fig. 5.4, 5.5, and 5.6 for CHOL, DLiPC, and DPPC, respectively. Four CHOL hot spots are repeated in every type of system in the intracellular leaflet of the bilayer: first between the helices 1 and 8, second on the surface of the helices 1 and 2, third near the helix 4, and fourth between the helices 3 and 4. In systems O1-O5 there is a CHOL hot spot on the surface of the helices 2 and 4 as well that is the location of the CHOL binding site found in the study of the β_2AR crystal structure [18]. The same CHOL hot spot, although a weaker one, is visible in the intracellular density map averaged over systems R1-R10, but not in the density map of the systems D1-D5.

The CHOL hot spots on the intracellular leaflet are all located on one side of β_2AR, whereas the hot spots on the extracellular leaflet roughly locate on the opposite side of the protein. The surface of the helices 5, 6, and 7 on the extracellular leaflet are occupied by CHOL in each type of system. Also a hot spot on the surface of the helix 4 can be seen in systems O1-O5, which is weaker in systems R1-R10

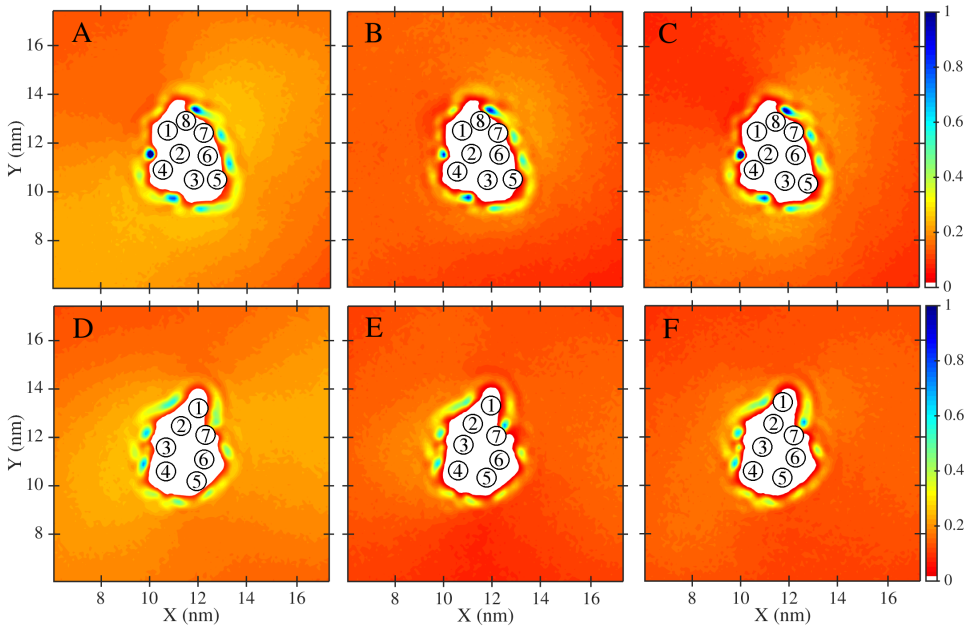


Figure 5.5: (A-C) The number density of DLiPC around β_2AR on the intracellular leaflet of the bilayer and (D-E) on the extracellular leaflet of the bilayer averaged over systems (A,D) R1-R10 (B,E), O1-O5, and (C,F) D1-D5.

and D1-D5. The CHOL hot spots on both sides of the bilayer are most prominent in the systems O1-O5, probably due to the fact that in these systems the protein is initially placed in the L_o phase enriched with CHOL. However, the hot spots are visible in all systems, irrespective of the initial configuration, and only small differences can be seen. The cholesterol hot spots close to the helices 1, 2, and 4 on the intracellular leaflet correspond to the location of cholesterol observed in the β_2AR crystal structure [16, 18], whereas the hot spots close to the helices 5, 6, and 7 in the extracellular leaflet are in agreement with cholesterol binding sites suggested for adenosine receptor A2A receptor, a member of GPCRs [193].

The spatial number density maps of DLiPC do not show such clear spots on the protein surface as the density maps of CHOL, but instead show almost a shell like distribution of DLiPC around β_2AR. In each type of systems, the shell gets slightly fainter on the intracellular surface of the helices 1 and 8. Similar density of DLiPC can be seen in each system.

The spatial distribution of DPPC does not show any remarkable hot spots on any side of the membrane, in fact, no DPPC molecules seem to locate near the β_2AR surface. There are minor visible spots in the density maps although none of them are as prominent as the hot spots of DLiPC and CHOL. In all systems there seems to be a small region preferred by DPPC on the protein surface, located on the

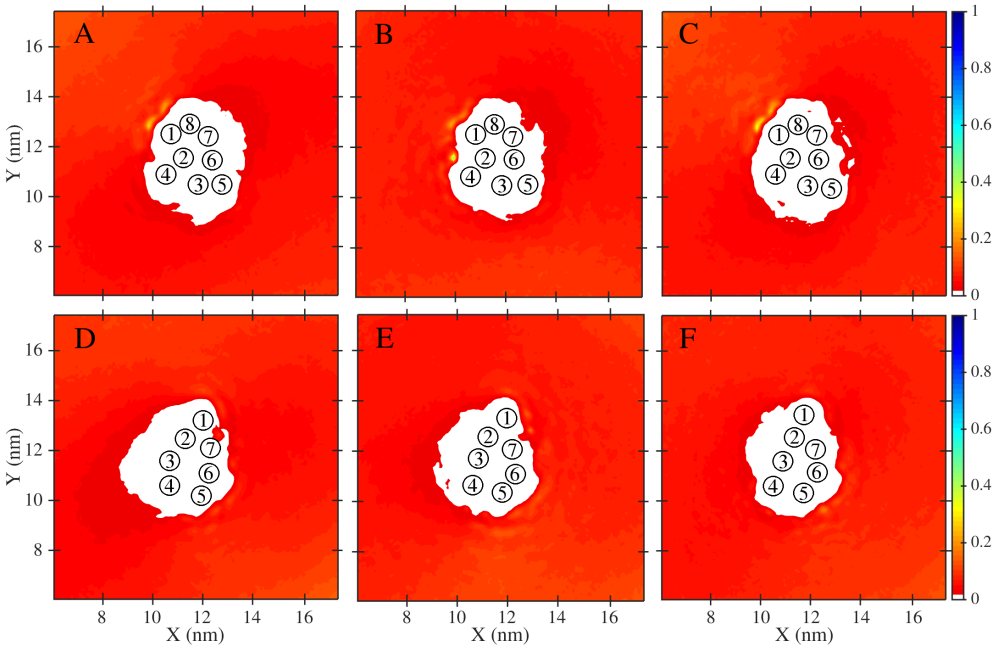


Figure 5.6: (A-C) The number density of DPPC around the β_2AR on the intracellular leaflet of the bilayer and (D-E) on the extracellular leaflet of the bilayer averaged over systems (A,D) R1-R10 (B,E), O1-O5, and (C,F) D1-D5.

intracellular surface of the helices 1 and 8. The region is consistent with the fainter area in the shell of DLiPC and with the location of the two most prominent CHOL hot spots on the intracellular leaflet.

In Fig. 5.7, 5.8, and 5.9 the lipid occupancy times per protein residue are averaged over systems with a similar starting configuration and plotted on the protein surface. There are several regions on the protein surface in contact with CHOL more than 60 % of time, including the intracellular part of the helices 1, 3, 4, 5, and 7, and the extracellular parts of the helices 2, 6, and 7. Most of these regions can also be seen as hot spots on the spatial density maps. In accordance with the density maps, the β_2AR surface is covered with regions where DLiPCs are staying more than 60 % of time, but in contrast only a few regions are in contact with DPPCs. The protein parts having contacts with DPPC are helix 6, extracellular part of the helix 2, and the intracellular parts of the helices 3 and 5. These regions are in correspondence to regions with high CHOL contact, especially helix 6.

The average lipid occupancy time per transmembrane part of each helix is shown in Fig. 5.10 averaged over each type of system. Each lipid is in contact with DLiPC the most, and the least with DPPC. On average the transmembrane residues of the helix 2 are almost 60 % of time in contact with DLiPC and almost 30 % of time

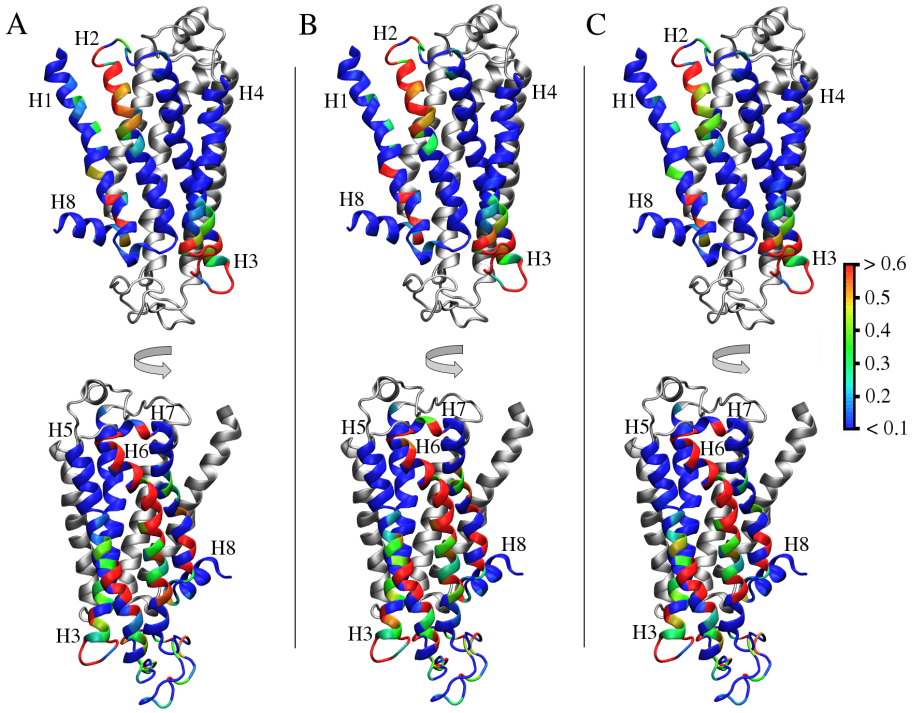


Figure 5.7: The cholesterol occupancy time on the β_2AR surface averaged over systems (A) R1-R10, (B) O1-O5, and (C) D1-D5.

with CHOL. The regions in which certain lipids tend to stay on the protein surface differ only slightly, indicating that irrespective of the initial location of the protein in the membrane, certain lipid types always localize to same regions on the protein surface. In this case the protein surface is occupied mostly by unsaturated DLiPC lipids and cholesterol.

For rhodopsin and opsin, other members of the GPCR superfamily, the polyunsaturated lipid chains were found as well packed against the surface of the protein [21]. Calculation of density maps and radial distribution of lipids yielded consistent results. In addition to the shell of polyunsaturated lipids surrounding the receptors, also few cholesterol hot spots were found. Polyunsaturated lipid chains have been found in high concentrations on the rhodopsin surface as well in other studies involving experiments and simulations [194–196]. This preference was suggested to be entropically driven [196].

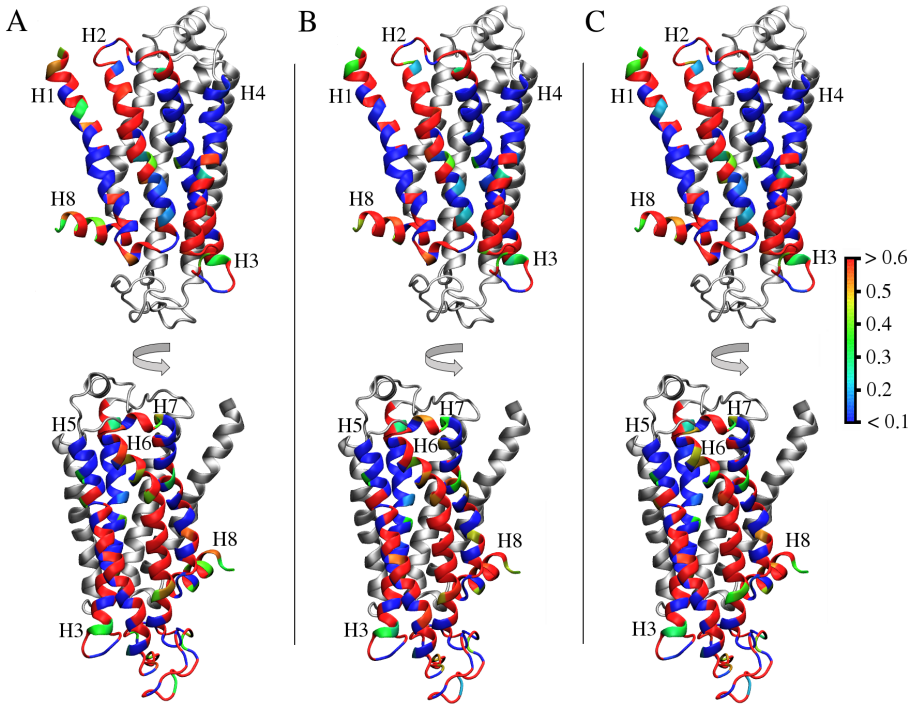


Figure 5.8: The DLiPC occupancy time on the β_2AR surface averaged over systems (A) R1-R10, (B) O1-O5, and (C) D1-D5.

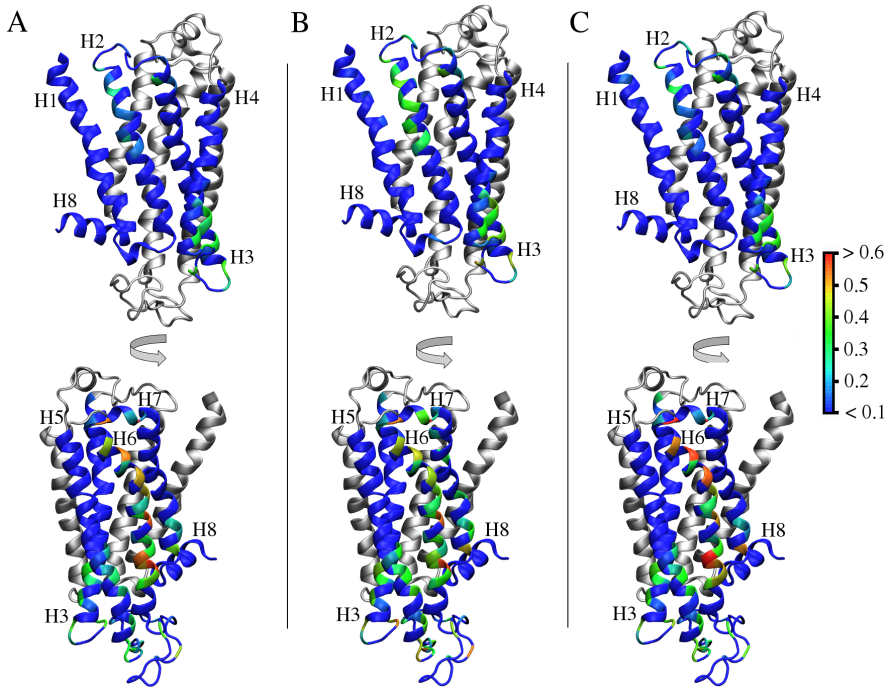


Figure 5.9: The DPPC occupancy time on the β_2AR surface averaged over systems (A) R1-R10, (B) O1-O5, and (C) D1-D5.

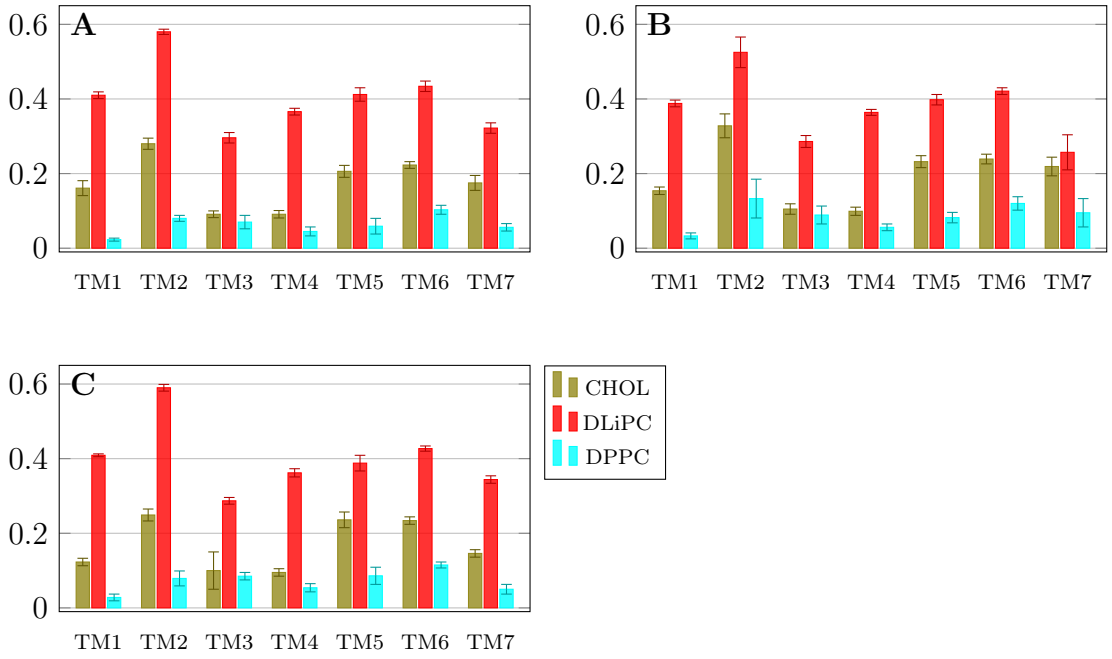


Figure 5.10: The lipid occupancy time per TM helix averaged over systems (A) R1-R10 (B) O1-O5, and (C) D1-D5.

5.4 Similar Phase Behaviour of β_2AR with Different Cholesterol Parameters

In the current Thesis, to check the validity of our results, we used two different cholesterol parameters, the one of Melo, Ingólfsson, and Marrink [172] and the one of Daily et al. [171]. The order parameters (Fig. 5.11(A-D)) and the diffusion coefficients (Fig. 5.11(E)) of the two coarse-grained PC lipids are quite different due to the different properties of the two phases. The diffusion coefficient of DLiPC varies between 4.4×10^{-8} cm^2/s and 5.1×10^{-8} cm^2/s in the systems R1-R10, O1-O5, and D1-D5 using the Martini CHOL parameter of Melo, Ingólfsson, and Marrink, and in the systems RC1 and RC2 with the CHOL parameter of Daily et al., the coefficient is $6.4\cdot10^{-8}$ cm^2/s. The diffusion coefficients of DPPC are between $1.6\cdot10^{-8}$ cm^2/s and $1.7\cdot10^{-8}$ cm^2/s in the systems R1-R10, O1-O5, and D1-D5, and $0.8\cdot10^{-8}$ cm^2/s in the systems RC1-RC2. The diffusion of lipids in the L_o phase is much slower than in the L_d phase.

The diffusion of lipids in the L_o phase is slower with the CHOL parameter of Daily et al. than with the one of Melo, Ingólfsson, and Marrink. The L_o phase is also more ordered, similar to the old Martini cholesterol [158,177,178]. The lipids in the L_d phase have a higher diffusion coefficient in systems RC1-RC2, and are thus

moving faster than with the CHOL parameter of Melo, Ingólfsson, and Marrink, but the order of the L_d phase is similar. A higher amount of cholesterol is located in the L_o phase than in the L_d phase, which causes a bigger influence in the order of the L_o phase comprising lipids when the cholesterol parameter is changed.

Even though the properties of the L_o and L_d phases differ when changing the CHOL parameter, the interaction of lipids with β_2AR is similar with both CHOL parameters used. The final RDF of lipids from the protein in the system RC2 (Fig. 5.11(F)) shows a similar distribution, and the number density of lipids gives a similar distribution of hot spots around the protein as the ones with the CHOL parameter of Melo, Ingólfsson, and Marrink. Also the lipid occupancy time averaged per transmembrane helix, shown in Fig. 5.11(G), gives similar preferences of lipids for certain helices of β_2AR as with the first parameter.

The CHOL parameter of Daily et al. has been developed to better match the properties of cholesterol in atomistic simulations. Yet, the effect of this cholesterol parameter in the behavior of the phase segregated membrane is similar to the effect of the previous Martini coarse-grained cholesterol parameter [158], that has been reported to cause the L_o phase to be too ordered [177, 178]. The Martini CHOL parameter of Melo, Ingólfsson, and Marrink used in 20 systems in this study better characterizes the phase behavior of the lipids than the parameter of Daily et al. that was used in two of the systems. Both of the parameters, however, exhibit similar behavior between the lipids and β_2AR.

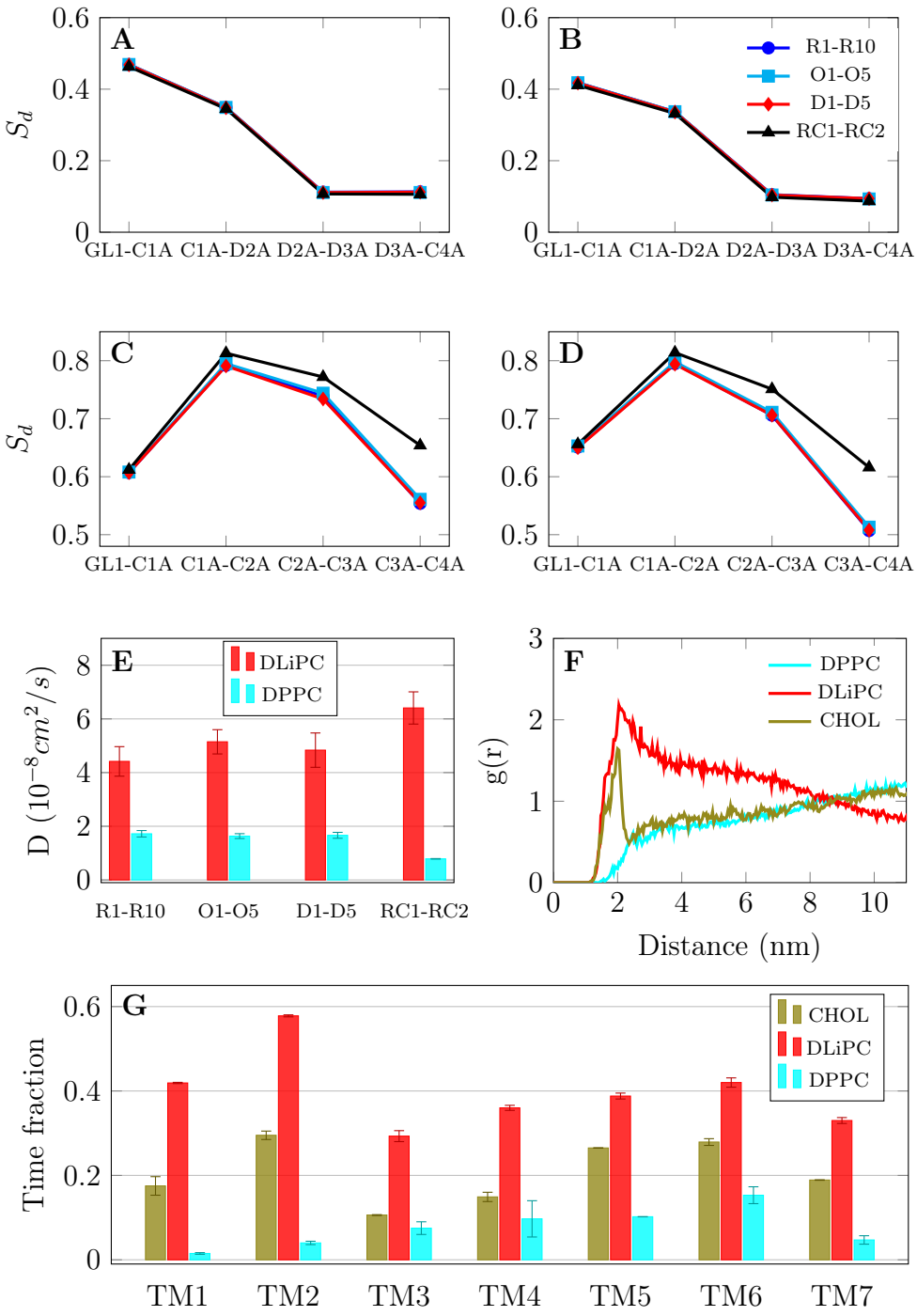


Figure 5.11: (A-D) The order parameters of the Martini coarse-grained PC lipid chains (A) DLiPC chain 1, (B) DLiPC chain 2, (C) DPPC chain 1, and (D) DPPC chain 2, averaged over the systems R1-R10, O1-O5, D1-D5, and RC1-RC2. Errors are less than 0.01. (E) The diffusion coefficients of the PC lipids fitted to the MSD curve averaged over the systems R1-R10, O1-O5, D1-D5, and RC1-RC2. (F) The radial distribution of lipids from the protein during the last 2 μs simulation in system RC2. (G) The lipid occupancy time per TM helix averaged over systems RC1-RC2.

5.5 β_2AR Prefers a Thinner and Less Ordered Local Membrane

In Fig. 5.12(A-D) the two-dimensional thickness averaged over each type of system is shown. The membrane thickness surrounding β_2AR is mainly between 39 and 40 Å in each type of system. The thickness corresponds to the membrane thickness in the L_d phase, while the L_o phase is thicker, up to 44 Å. Yet there are two regions repeated in the 2D thickness plot of each system type in which the bilayer is thicker than in the L_d phase, almost 41 to 42 Å. These regions are located close to the helices 1 and 8, and the helices 3, 4, and 5. Cholesterol number density showed hot spots present in this area, which might induce the membrane thickness at these locations. The hydrophobic thickness of the membrane in the L_o/L_d boundary might match the hydrophobic parts of the protein and this could be one of the reasons for the protein to localize to the phase boundary. It seems that the protein creates a special thickness environment around it, with slightly varying thicknesses.

The lipid chain order parameter averaged over all the systems is shown in Fig. 5.12(E-F). The protein seems to prefer a less ordered membrane around it. The two thicker regions surrounding the protein, visible in the 2D thickness plots, are not that clearly repeated in the membrane order plots. It has been postulated that the lipid type surrounding the protein might define its affinity for domains consisting of lipids compatible with the lipids surrounding the protein [82]. Since the protein is surrounded by a more disordered membrane, the protein might therefore end up to the side of the L_d domain in the phase boundary. However, it is obvious that the β_2AR prefers a thinner membrane and less ordered lipid chains around it than in the L_o phase, and seems to create a special environment of slightly varying properties around it.

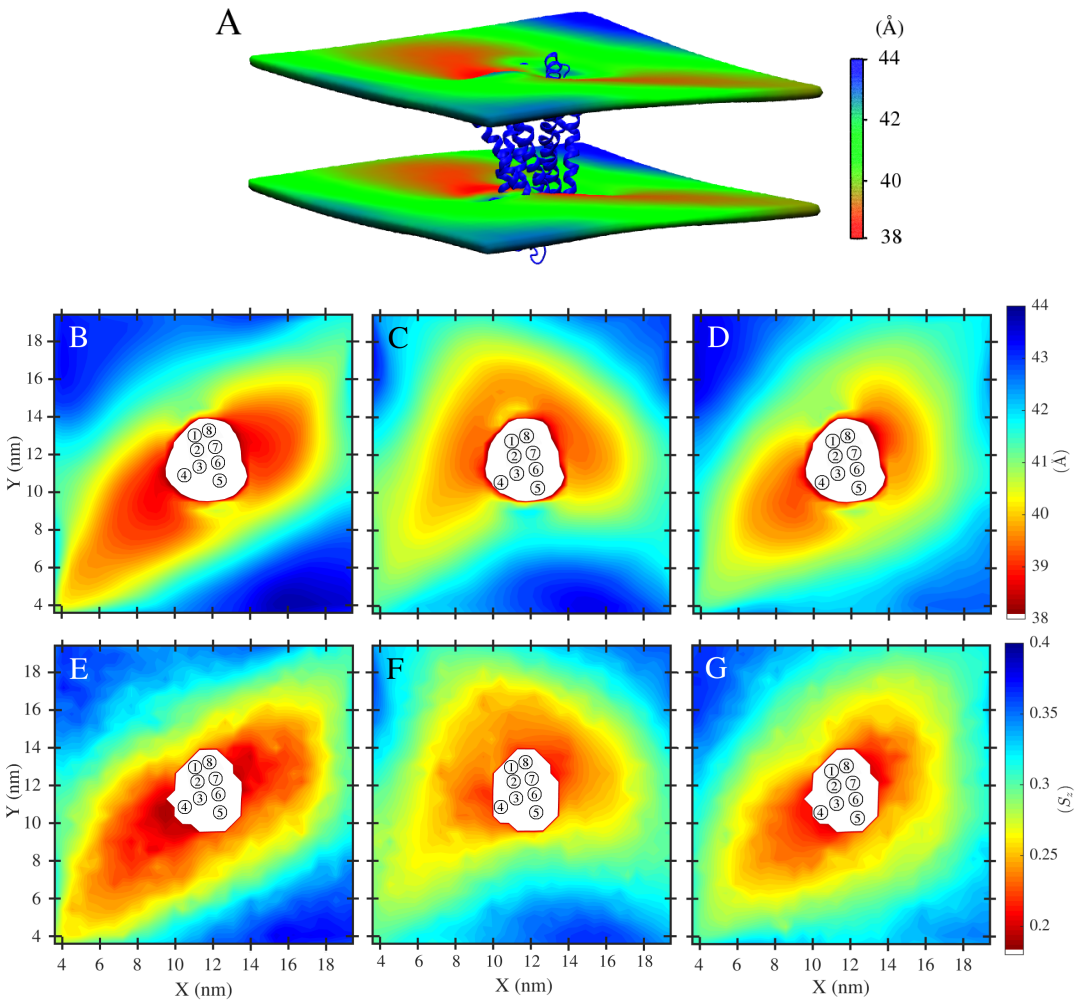


Figure 5.12: (A) A 3D distribution of the bilayer thickness averaged over systems R1-R10. (B-D) 2D thickness and (E-G) lipid order parameter averaged over systems (B,E) R1-R10, (C,F) O1-O5, and (D,G) D1-D5.

5.6 Atomistic Systems Show Similar Properties as Coarse-Grained Systems

The systems B1 and B2 were backmapped from a coarse-grained to an all-atom description and their simulations were started from configurations in which the protein was in the L_o/L_d phase boundary (system B1) or completely in the L_d phase (system B2). In Fig. 5.13, the initial configurations of the systems B1 and B2 are shown along with the configurations after 2 μs simulation. In the system B1, the protein is in the phase boundary, and during the simulation the few saturated DPPC molecules on its surface are replaced by unsaturated DOPC and CHOL molecules,

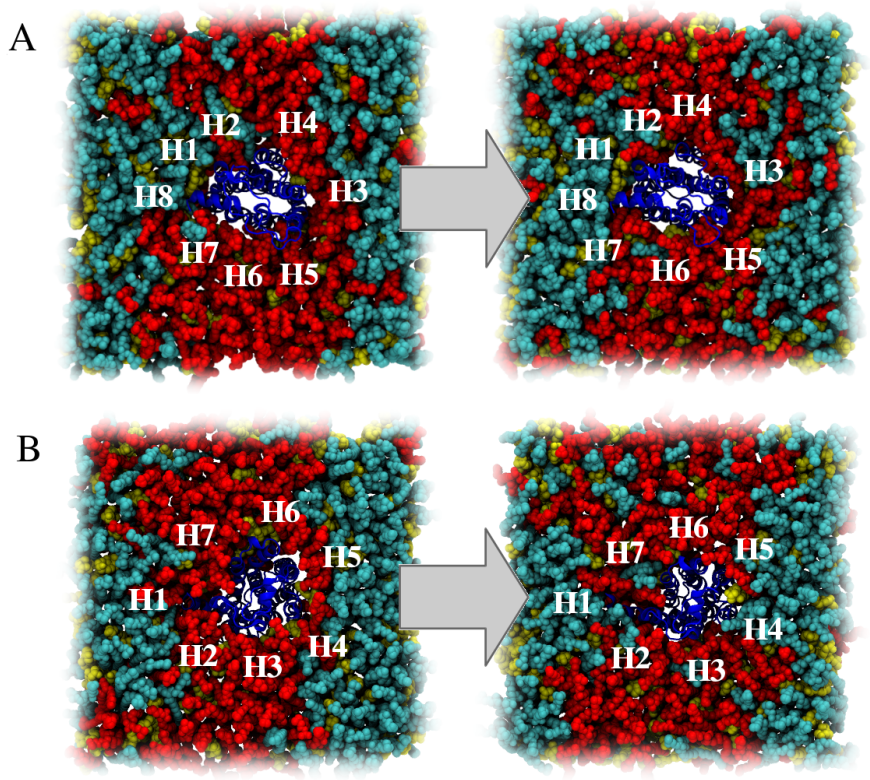


Figure 5.13: Snapshots from the beginning and after 2 μs simulation of (A) the intracellular side of the system B1 and (B) the extracellular side of the system B2. The color codes are similar to Fig. 4.1 and Fig. 4.2. The protein helices are marked with white labels next to the helices.

although saturated lipids remain close to the helices 1, 7, and 8. The system B2 is completely surrounded by the L_d phase in the beginning of the simulation, and remains there, but in addition some saturated lipids move close to the helices 1, 2, 4, and 7. In these locations the membrane was found slightly thicker compared to the total membrane surrounding the protein in the coarse-grained simulations (Fig. 5.12).

The fraction of lipid-lipid and protein-lipid contacts in the systems B1 and B2 are plotted in Fig. 5.14(A-B). The lipid-lipid and protein-lipid contact fraction curves are similar to the ones obtained in coarse-grained systems, although it seems that especially the protein-lipid contact fraction in system B2 is not stabilized within the 2 μs simulation. However, the simulations are very short compared to the coarse-grained simulations, and the contact fraction curves are fluctuating within reasonable limits if compared to the curves in coarse-grained simulations during the same time interval.

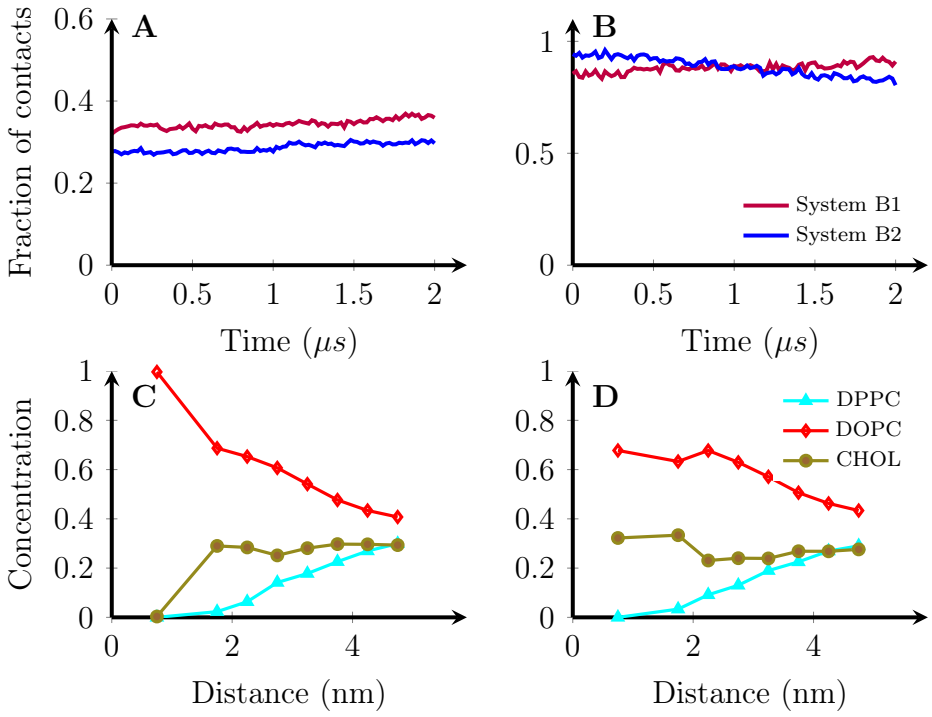


Figure 5.14: (A) The fraction of lipid-lipid contacts, and (B) the fraction of protein-lipid contacts in the atomistic systems B1 and B2. The lipid composition versus distance from the protein calculated from the last 1 μs trajectory of the systems (C) B1 and (D) B2.

The lipid composition close to the protein (Fig. 5.14(C-D)) is similar to the coarse-grained systems at a distance of 1.5-2.5 nm. The unsaturated lipids are the dominating ones along with a faint peak in the cholesterol concentration, while the saturated lipids are excluded from the protein surface. The cholesterol concentration remains similar to the coarse-grained systems, around 30 %, yet the unsaturated DOPC lipid concentration is higher, close to 70 %, whereas the saturated DPPC concentration stays very low, less than 10 %.

It should be noted that the analysis results from the two atomistic systems are only single repetitions, whereas the results from the coarse-grained systems have been averaged over many systems with longer simulation times. Additionally, the protein partitioning from similar starting configurations as in the coarse-grained systems is quite challenging with the current atom-scale computing limits, for which reason the systems were started near the assumed protein equilibrium location. However, the 2 μs simulation of atomistic systems is enough for the protein to slightly stabilize its surroundings, and no dramatic changes are seen, which indicates that the obtained environment is appropriate for the protein also in the atomistic model.

6. CONCLUSIONS

The aim of this Thesis was to find out the preference of the β_2–adrenergic receptor, a characteristic member of the GPCRs, for different membrane domains in a laterally phase segregated membrane. Many components important in the cAMP pathway, such as GPCRs and G proteins, have been suggested to localize to caveolae and rafts that would enable rapid activation of signaling pathways [10, 54, 58, 132, 133]. β_1AR and β_2AR have been found to localize in caveolae, but whereas β_1AR continually stays there, β_2AR was found to translocate out of the caveolae upon activation [11]. The β_2AR partitioning into domains has not been studied either by simulations or membrane vesicle studies, although such studies for rhodopsin show partitioning to the L_d domain [19, 20]. Additionally, some simulation studies of rhodopsin and opsin, other GPCRs, show preference for unsaturated lipid chains on the receptor surface [21]. These studies would support the partitioning of β_2AR, a receptor reminiscent to rhodopsin, to the L_d phase. However, β_2AR is also widely associated to cholesterol [16–18], which on the other hand could lead to the association of the receptor to the cholesterol-enriched L_o domain.

The problem was approached by computational molecular dynamics simulations using both the coarse-grained Martini force field and the all-atom OPLS force field for the simulated lipids and the receptor protein. In total, 22 coarse-grained systems were built in which the inactive β_2AR was placed in different locations; ten systems with the protein initially in a random lipid distribution, and ten systems with a completely phase segregated bilayer in which the protein was placed in the L_o phase (five systems) or in the L_d phase (five systems), and two systems in which another coarse-grained cholesterol model was tested. The two atomistic systems were backmapped from the coarse-grained systems in order to validate the results obtained by the coarse-grained simulations; in one system the protein was completely in the L_d phase surrounded by unsaturated lipids, and in the other system the protein was at the boundary of the two phases.

The results obtained from the simulations all supported the idea of β_2AR partitioning to the L_d domain, however instead of being completely embedded to the domain, the protein was found to reside at the boundary of the two phases, though

most of the time surrounded by unsaturated lipids. Both concentration and radial distribution function of lipids showed a peak close to the protein in the distribution of the unsaturated lipids and cholesterol, while saturated lipids were in all cases staying further away from the protein. Notable is that the results obtained from all simulations were similar, irrespective of the location in which the protein was initially placed, which implies that the protein found its equilibrium location during the simulation time in each system.

The density maps of lipids, the local membrane thickness, and the local lipid chain order parameter all indicated that β_2AR creates a special lipid environment around it. The local membrane around the protein comprised a shell of unsaturated lipids and hot spots of cholesterol. The surrounding lipid environment was very similar to the ones obtained for rhodopsin and opsin [21, 194–196]. The protein was also residing in a less thicker and less ordered membrane environment than the L_o phase. The membrane surrounding the protein was not completely uniform in thickness, but included a few locations with higher thickness. These locations were in accordance with the cholesterol hot spots, which might lead to the thickening of the membrane.

The results obtained from the atomistic systems were in accordance with the ones obtained from the coarse-grained systems. Due to computational limits, the atomistic systems were started from the configuration in which the protein was close to the assumed equilibrium location. In order to better justify the surrounding environment preferred by β_2AR, another atomistic system will be built in which the protein will be placed in a random lipid distribution. Even though the time scale of lipid phase separation is unreachable, the lipid-protein environment should stabilize within a few microseconds.

There might be several reasons for the β_2AR partitioning into the L_o/L_d phase boundary. Transmembrane proteins can alter the surrounding membrane properties up to 3-4 nm distance away from the protein surface [83]. For lower free energy cost, the membrane and the protein always tend to organize in such a way that the hydrophobic thickness of the membrane matches with the hydrophobic regions of the protein [84–88]. This might be one factor affecting the protein location, however it has been shown that the hydrophobic thickness can not be the only driving force in the protein segregation to the domains [146, 197]. Another suggestion would be the membrane diffusion rate matching the protein diffusion in that area. It has also been suggested that the lipids surrounding the transmembrane protein might define its affinity for domains composed of lipids compatible with ones surrounding the

protein [82]. This theory would explain the protein preference for the L_d domain composed of unsaturated lipids that create the shell around the protein. On the other hand, the protein does not completely stay in the L_d domain, but at the boundary of the two domains. The results indicate that several cholesterol molecules also reside on the protein surface and that the cholesterol concentration around the protein is close to the concentration of the L_o phase, which might drag the protein towards the L_o phase enriched with cholesterol. The affinity for two domains might result in the protein staying in the boundary of the two phases. The reason for the protein staying on the side of the L_d phase at the boundary might be in its preference for the unsaturated lipids chains [21, 194–196].

The results obtained in this study are consistent with the ones obtained for rhodopsin, which was found in simulations and giant vesicle studies to reside in the L_d phase [19, 20]. In addition to residing mostly in the L_d phase, this study is the first one to show that β_2AR creates its own environment at the boundary of the two phases. Meanwhile, the results are not consistent with the suggestion of the β_2AR residing in the L_o domain in an inactive state, at least with this molecular model [11]. There are other factors as well that might affect the protein partitioning, such as glycosylation and palmitoylation [90], which could be applied to see if the β_2AR preference for the surrounding lipids and for the location with respect to the phases would change. Since activity–dependent partitioning of β_2AR has been suggested [11], and also the active, Gs protein bound structure of β_2AR has been discovered, similar computational studies involving the active structure with and/or without the bound G protein could be carried out to investigate this hypothesis. Combining experimental studies of the β_2AR partitioning in for example phase segregated giant vesicles, as for rhodopsin [19], would bring new insights to the matter. In the future, along with increasing computational capacity, the time scales of receptor partitioning in a phase segregated membrane could even be reachable for atomic scale systems.

BIBLIOGRAPHY

[1] **Marsh, D.**, Cholesterol-induced fluid membrane domains: a compendium of lipid-raft ternary phase diagrams. *Biochimica et Biophysica Acta (BBA)-Biomembranes*, 1788(10): 2114–2123. 2009.

[2] **Ackerman, D.G. and Feigenson, G.W.**, Lipid bilayers: clusters, domains and phases. *Essays in Biochemistry*, 57: 33–42. 2015.

[3] **Veatch, S.L. and Keller, S.L.**, A closer look at the canonical 'raft mixture' in model membrane studies. *Biophysical Journal*, 84(1): 725–726. 2003.

[4] **Dietrich, C., Bagatolli, L., Volovyk, Z., Thompson, N., Levi, M., Jacobson, K., and Gratton, E.**, Lipid rafts reconstituted in model membranes. *Biophysical Journal*, 80(3): 1417–1428. 2001.

[5] **Gandhavadi, M., Allende, D., Vidal, A., Simon, S., and McIntosh, T.**, Structure, composition, and peptide binding properties of detergent soluble bilayers and detergent resistant rafts. *Biophysical Journal*, 82(3): 1469–1482. 2002.

[6] **Simons, K. and Ikonen, E.**, Functional rafts in cell membranes. *Nature*, 387(6633): 569–572. 1997.

[7] **Takeda, S., Kadowaki, S., Haga, T., Takaesu, H., and Mitaku, S.**, Identification of G protein-coupled receptor genes from the human genome sequence. *FEBS Letters*, 520(1): 97–101. 2002.

[8] **Cheng, Z., Garvin, D., Paguio, A., Stecha, P., Wood, K., and Fan, F.**, Luciferase reporter assay system for deciphering GPCR pathways. *Current Chemical Genomics*, 4(1). 2010.

[9] **Rasmussen, S.G., DeVree, B.T., Zou, Y., Kruse, A.C., Chung, K.Y., Kobilka, T.S., Thian, F.S., Chae, P.S., Pardon, E., Calinski, D., et al.**, Crystal structure of the β 2-adrenergic receptor-Gs protein complex. *Nature*, 477(7366): 549–555. 2011.

[10] **Patel, H., Murray, F., and Insel, P.**, G-protein-coupled receptor-signaling components in membrane raft and caveolae microdomains. In *Protein-Protein Interactions as New Drug Targets*, pp. 167–184, Springer. 2008.

[11] **Rybin, V.O., Xu, X., Lisanti, M.P., and Steinberg, S.F.**, Differential targeting of β-adrenergic receptor subtypes and adenylyl cyclase to cardiomyocyte caveolae A mechanism to functionally regulate the cAMP signaling pathway. *Journal of Biological Chemistry*, 275(52): 41447–41457. 2000.

[12] **Perez, J.B., Segura, J.M., Abankwa, D., Piguet, J., Martinez, K.L., and Vogel, H.**, Monitoring the diffusion of single heterotrimeric G proteins in supported cell-membrane sheets reveals their partitioning into microdomains. *Journal of Molecular Biology*, 363(5): 918–930. 2006.

[13] **Holst, B., Hastrup, H., Raffetseder, U., Martini, L., and Schwartz, T.W.**, Two active molecular phenotypes of the tachykinin NK1 receptor revealed by G-protein fusions and mutagenesis. *Journal of Biological Chemistry*, 276(23): 19793–19799. 2001.

[14] **Neubig, R.**, Membrane organization in G-protein mechanisms. *The FASEB Journal*, 8(12): 939–946. 1994.

[15] **Rebois, R.V. and Hébert, T.E.**, Protein complexes involved in heptahelical receptor-mediated signal transduction. *Receptors and Channels*, 9(3): 169–194. 2003.

[16] **Cherezov, V., Rosenbaum, D.M., Hanson, M.A., Rasmussen, S.G., Thian, F.S., Kobilka, T.S., Choi, H.J., Kuhn, P., Weis, W.I., Kobilka, B.K., et al.**, High-resolution crystal structure of an engineered human β2-adrenergic G protein–coupled receptor. *Science*, 318(5854): 1258–1265. 2007.

[17] **Zocher, M., Zhang, C., Rasmussen, S.G., Kobilka, B.K., and Müller, D.J.**, Cholesterol increases kinetic, energetic, and mechanical stability of the human β2-adrenergic receptor. *Proceedings of the National Academy of Sciences*, 109(50): E3463–E3472. 2012.

[18] **Hanson, M.A., Cherezov, V., Griffith, M.T., Roth, C.B., Jaakola, V.P., Chien, E.Y., Velasquez, J., Kuhn, P., and Stevens, R.C.**, A specific cholesterol binding site is established by the 2.8 Å structure of the human β2-adrenergic receptor. *Structure*, 16(6): 897–905. 2008.

[19] **Kahya, N., Brown, D.A., and Schwille, P.**, Raft partitioning and dynamic behavior of human placental alkaline phosphatase in giant unilamellar vesicles. *Biochemistry*, 44(20): 7479–7489. 2005.

[20] **Domański, J., Marrink, S.J., and Schäfer, L.V.**, Transmembrane helices can induce domain formation in crowded model membranes. *Biochimica et Biophysica Acta (BBA)-Biomembranes*, 1818(4): 984–994. 2012.

[21] **Horn, J.N., Kao, T.C., and Grossfield, A.**, Coarse-grained molecular dynamics provides insight into the interactions of lipids and cholesterol with rhodopsin. In *G Protein-Coupled Receptors - Modeling and Simulation*, pp. 75–94, Springer. 2014.

[22] **Mllligan, G., Svoboda, P., and Brown, C.M.**, Why are there so many adrenoceptor subtypes? *Biochemical Pharmacology*, 48(6): 1059–1071. 1994.

[23] **Billman, G.E., Castillo, L.C., Hensley, J., Hohl, C.M., and Altschuld, R.A.**, β2-adrenergic receptor antagonists protect against ventricular fibrillation in vivo and in vitro evidence for enhanced sensitivity to β2-adrenergic stimulation in animals susceptible to sudden death. *Circulation*, 96(6): 1914–1922. 1997.

[24] **Cazzola, M., Calzetta, L., and Matera, M.G.**, β2-adrenoceptor agonists: current and future direction. *British Journal of Pharmacology*, 163(1): 4–17. 2011.

[25] **Singer, S.J. and Nicolson, G.L.**, The fluid mosaic model of the structure of cell membranes. *Science*, 175(4023): 720–731. 1972.

[26] **Alberts, B., Bray, D., Hopkin, K., Johnson, A., Lewis, J., Raff, M., Roberts, K., and Walter, P.**, *Essential Cell Biology*. Garland Science, 3. edition. 2010, 121–132,363–548 pp.

[27] **Van Meer, G., Voelker, D.R., and Feigenson, G.W.**, Membrane lipids: where they are and how they behave. *Nature Reviews Molecular Cell Biology*, 9(2): 112–124. 2008.

[28] **Luckey, M.**, *Membrane Structural Biology: with Biochemical and Biophysical Foundations*. Cambridge University Press, 2. edition. 2014, 21–25 pp.

[29] **Dwek, R.A.**, Glycobiology: toward understanding the function of sugars. *Chemical Reviews*, 96(2): 683–720. 1996.

[30] **Karnovsky, M.J., Kleinfeld, A.M., Hoover, R.L., and Klausner, R.D.**, The concept of lipid domains in membranes. *The Journal of Cell Biology*, 94(1): 1–6. 1982.

[31] **Ipsen, J.H., Karlström, G., Mourtisen, O., Wennerström, H., and Zuckermann, M.**, Phase equilibria in the phosphatidylcholine-cholesterol system. *Biochimica et Biophysica Acta (BBA)-Biomembranes*, 905(1): 162–172. 1987.

[32] **Lentz, B.R., Barrow, D.A., and Hoechli, M.**, Cholesterol-phosphatidylcholine interactions in multilamellar vesicles. *Biochemistry*, 19(9): 1943–1954. 1980.

[33] **Simons, K. and Van Meer, G.**, Lipid sorting in epithelial cells. *Biochemistry*, 27(17): 6197–6202. 1988.

[34] **Edidin, M.**, The state of lipid rafts: from model membranes to cells. *Annual Review of Biophysics and Biomolecular Structure*, 32(1): 257–283. 2003.

[35] **Koumanov, K.S., Wolf, C., and Quinn, P.J.**, Lipid composition of membrane domains. In *Membrane Dynamics and Domains*, pp. 153–163, Springer. 2004.

[36] **Nicolson, G.L.**, The fluid-mosaic model of membrane structure: still relevant to understanding the structure, function and dynamics of biological membranes after more than 40 years. *Biochimica et Biophysica Acta (BBA)-Biomembranes*, 1838(6): 1451–1466. 2014.

[37] **Owen, D.M., Magenau, A., Williamson, D., and Gaus, K.**, The lipid raft hypothesis revisited–New insights on raft composition and function from super-resolution fluorescence microscopy. *Bioessays*, 34(9): 739–747. 2012.

[38] **Pike, L.J.**, Lipid rafts bringing order to chaos. *Journal of Lipid Research*, 44(4): 655–667. 2003.

[39] **Pralle, A., Keller, P., Florin, E.L., Simons, K., and Hörber, J.H.**, Sphingolipid–cholesterol rafts diffuse as small entities in the plasma membrane of mammalian cells. *The Journal of Cell Biology*, 148(5): 997–1008. 2000.

[40] **Simons, K. and Sampaio, J.L.**, Membrane organization and lipid rafts. *Cold Spring Harbor Perspectives in Biology*, 3(10): a004697. 2011.

[41] **Lingwood, D. and Simons, K.**, Lipid rafts as a membrane-organizing principle. *Science*, 327(5961): 46–50. 2010.

[42] **Yeagle, P.L.**, Cholesterol and the cell membrane. *Biochimica et Biophysica Acta (BBA)-Reviews on Biomembranes*, 822(3): 267–287. 1985.

[43] **Simons, K. and Vaz, W.L.**, Model systems, lipid rafts, and cell membranes 1. *Annual Review of Biophysics and Biomolecular Structure*, 33: 269–295. 2004.

[44] **Veatch, S.L. and Keller, S.L.**, Miscibility phase diagrams of giant vesicles containing sphingomyelin. *Physical Review Letters*, 94(14): 148101. 2005.

[45] **Brown, D.A.**, Lipid rafts, detergent-resistant membranes, and raft targeting signals. *Physiology*, 21(6): 430–439. 2006.

[46] **Murata, M., Peränen, J., Schreiner, R., Wieland, F., Kurzchalia, T.V., and Simons, K.**, VIP21/caveolin is a cholesterol-binding protein. *Proceedings of the National Academy of Sciences*, 92(22): 10339–10343. 1995.

[47] **Palade, G.E.**, Fine structure of blood capillaries. *J Appl Phys*, 24(1): 1424–1436. 1953.

[48] **Yamada, E.**, The fine structure of the gall bladder epithelium of the mouse. *The Journal of Biophysical and Biochemical Cytology*, 1(5): 445. 1955.

[49] **Fra, A.M., Williamson, E., Simons, K., and Parton, R.G.**, De novo formation of caveolae in lymphocytes by expression of VIP21-caveolin. *Proceedings of the National Academy of Sciences*, 92(19): 8655–8659. 1995.

[50] **Le, P.U., Guay, G., Altschuler, Y., and Nabi, I.R.**, Caveolin-1 is a negative regulator of caveolae-mediated endocytosis to the endoplasmic reticulum. *Journal of Biological Chemistry*, 277(5): 3371–3379. 2002.

[51] **Li, S., Song, K.S., and Lisanti, M.P.**, Expression and characterization of recombinant caveolin. Purification by polyhistidine tagging and cholesterol-dependent incorporation into defined lipid membranes. *Journal of Biological Chemistry*, 271(1): 568–573. 1996.

[52] **Simons, K. and Toomre, D.**, Lipid rafts and signal transduction. *Nature Reviews Molecular Cell Biology*, 1(1): 31–39. 2000.

[53] **Smart, E.J., Graf, G.A., McNiven, M.A., Sessa, W.C., Engelman, J.A., Scherer, P.E., Okamoto, T., and Lisanti, M.P.**, Caveolins, liquid-ordered domains, and signal transduction. *Molecular and Cellular Biology*, 19(11): 7289–7304. 1999.

[54] **Cohen, A.W., Hnasko, R., Schubert, W., and Lisanti, M.P.**, Role of caveolae and caveolins in health and disease. *Physiological Reviews*, 84(4): 1341–1379. 2004.

[55] **Liu, P., Rudick, M., and Anderson, R.G.**, Multiple functions of caveolin-1. *Journal of Biological Chemistry*, 277(44): 41295–41298. 2002.

[56] **Morris, R., Cox, H., Mombelli, E., and Quinn, P.J.**, Rafts, little caves and large potholes: how lipid structure interacts with membrane proteins to create functionally diverse membrane environments. In *Membrane Dynamics and Domains*, pp. 35–118, Springer. 2004.

[57] **Parton, R.G. and Simons, K.**, The multiple faces of caveolae. *Nature Reviews Molecular Cell Biology*, 8(3): 185–194. 2007.

[58] **van Deurs, B., Roepstorff, K., Hommelgaard, A.M., and Sandvig, K.**, Caveolae: anchored, multifunctional platforms in the lipid ocean. *Trends in Cell Biology*, 13(2): 92–100. 2003.

[59] **Veatch, S., Polozov, I., Gawrisch, K., and Keller, S.**, Liquid domains in vesicles investigated by NMR and fluorescence microscopy. *Biophysical Journal*, 86(5): 2910–2922. 2004.

[60] **Hancock, J.F.**, Lipid rafts: contentious only from simplistic standpoints. *Nature Reviews Molecular Cell Biology*, 7(6): 456–462. 2006.

[61] **Leslie, M.**, Do lipid rafts exist? *Science*, 334(6059): 1046–1047. 2011.

[62] **Pike, L.J.**, Rafts defined: a report on the Keystone Symposium on lipid rafts and cell function. *Journal of Lipid Research*, 47(7): 1597–1598. 2006.

[63] **Zurzolo, C., Van Meer, G., and Mayor, S.**, The order of rafts. *EMBO Reports*, 4(12): 1117–1121. 2003.

[64] **Sevcsik, E. and Schütz, G.J.**, With or without rafts? Alternative views on cell membranes. *Bioessays*, 38(2): 129–139. 2016.

[65] **Simons, K. and Ehehalt, R.**, Cholesterol, lipid rafts, and disease. *The Journal of Clinical Investigation*, 110(5): 597–603. 2002.

[66] **Shogomori, H. and Brown, D.A.**, Use of detergents to study membrane rafts: the good, the bad, and the ugly. *Biological Chemistry*, 384(9): 1259–1263. 2003.

[67] **Heerklotz, H.**, Triton promotes domain formation in lipid raft mixtures. *Biophysical Journal*, 83(5): 2693–2701. 2002.

[68] **Schroeder, R., London, E., and Brown, D.**, Interactions between saturated acyl chains confer detergent resistance on lipids and glycosylphosphatidylinositol (GPI)-anchored proteins: GPI-anchored proteins in liposomes and cells show similar behavior. *Proceedings of the National Academy of Sciences*, 91(25): 12130–12134. 1994.

[69] **van Meer, G. and Simons, K.**, Viruses budding from either the apical or the basolateral plasma membrane domain of MDCK cells have unique phospholipid compositions. *The EMBO Journal*, 1(7): 847. 1982.

[70] **Harder, T., Scheiffele, P., Verkade, P., and Simons, K.**, Lipid domain structure of the plasma membrane revealed by patching of membrane components. *The Journal of Cell Biology*, 141(4): 929–942. 1998.

[71] **Yetukuri, L., Ekroos, K., Vidal-Puig, A., and Orešič, M.**, Informatics and computational strategies for the study of lipids. *Molecular BioSystems*, 4(2): 121–127. 2008.

[72] **Dupuy, A.D. and Engelman, D.M.**, Protein area occupancy at the center of the red blood cell membrane. *Proceedings of the National Academy of Sciences*, 105(8): 2848–2852. 2008.

[73] **Scheve, C.S., Gonzales, P.A., Momin, N., and Stachowiak, J.C.**, Steric pressure between membrane-bound proteins opposes lipid phase separation. *Journal of the American Chemical Society*, 135(4): 1185–1188. 2013.

[74] **Baumgart, T., Hammond, A.T., Sengupta, P., Hess, S.T., Holowka, D.A., Baird, B.A., and Webb, W.W.**, Large-scale fluid/fluid phase separation of proteins and lipids in giant plasma membrane vesicles. *Proceedings of the National Academy of Sciences*, 104(9): 3165–3170. 2007.

[75] **Ma, B., Shatsky, M., Wolfson, H.J., and Nussinov, R.**, Multiple diverse ligands binding at a single protein site: A matter of pre-existing populations. *Protein Science*, 11(2): 184–197. 2002.

[76] **Lee, A.G.**, How lipids affect the activities of integral membrane proteins. *Biochimica et Biophysica Acta (BBA)-Biomembranes*, 1666(1): 62–87. 2004.

[77] **Phillips, R., Ursell, T., Wiggins, P., and Sens, P.**, Emerging roles for lipids in shaping membrane-protein function. *Nature*, 459(7245): 379–385. 2009.

[78] **Laganowsky, A., Reading, E., Allison, T.M., Ulmschneider, M.B., Degiacomi, M.T., Baldwin, A.J., and Robinson, C.V.**, Membrane proteins bind lipids selectively to modulate their structure and function. *Nature*, 510(7503): 172. 2014.

[79] **Lee, A.G.**, Biological membranes: the importance of molecular detail. *Trends in Biochemical Sciences*, 36(9): 493–500. 2011.

[80] **Lee, A.**, Lipid–protein interactions in biological membranes: a structural perspective. *Biochimica et Biophysica Acta (BBA)-Biomembranes*, 1612(1): 1–40. 2003.

[81] **Marsh, D.**, Protein modulation of lipids, and vice-versa, in membranes. *Biochimica et Biophysica Acta (BBA)-Biomembranes*, 1778(7): 1545–1575. 2008.

[82] **Anderson, R.G. and Jacobson, K.**, A role for lipid shells in targeting proteins to caveolae, rafts, and other lipid domains. *Science*, 296(5574): 1821–1825. 2002.

[83] **Niemelä, P.S., Miettinen, M.S., Monticelli, L., Hammaren, H., Bjelkmar, P., Murtola, T., Lindahl, E., and Vattulainen, I.**, Membrane proteins diffuse as dynamic complexes with lipids. *Journal of the American Chemical Society*, 132(22): 7574–7575. 2010.

[84] **Dumas, F., Lebrun, M.C., and Tocanne, J.F.**, Is the protein/lipid hydrophobic matching principle relevant to membrane organization and functions? *FEBS Letters*, 458(3): 271–277. 1999.

[85] **Kusumi, A., Suzuki, K.G., Kasai, R.S., Ritchie, K., and Fujiwara, T.K.**, Hierarchical mesoscale domain organization of the plasma membrane. *Trends in Biochemical Sciences*, 36(11): 604–615. 2011.

[86] **Mohandas, N. and Evans, E.**, Mechanical properties of the red cell membrane in relation to molecular structure and genetic defects. *Annual Review of Biophysics and Biomolecular Structure*, 23(1): 787–818. 1994.

[87] **Mouritsen, O.G.**, Model answers to lipid membrane questions. *Cold Spring Harbor Perspectives in Biology*, 3(9): a004622. 2011.

[88] **Mouritsen, O.G.**, Lipids, curvature, and nano-medicine. *European Journal of Lipid Science and Technology*, 113(10): 1174–1187. 2011.

[89] **Gil, T., Sabra, M.C., Ipsen, J.H., and Mouritsen, O.G.**, Wetting and capillary condensation as means of protein organization in membranes. *Biophysical Journal*, 73(4): 1728. 1997.

[90] **Levental, I., Lingwood, D., Grzybek, M., Coskun, Ü., and Simons, K.**, Palmitoylation regulates raft affinity for the majority of integral raft proteins. *Proceedings of the National Academy of Sciences*, 107(51): 22050–22054. 2010.

[91] **Zacharias, D.A., Violin, J.D., Newton, A.C., and Tsien, R.Y.**, Partitioning of lipid-modified monomeric GFPs into membrane microdomains of live cells. *Science*, 296(5569): 913–916. 2002.

[92] **Carpino, P.A. and Goodwin, B.**, Diabetes area participation analysis: a review of companies and targets described in the 2008–2010 patent literature. *Expert Opinion on Therapeutic Patents*, 20(12): 1627–1651. 2010.

[93] **Parma, J., Duprez, L., Van Sande, J., Cochaux, P., Gervy, C., Mockel, J., Dumont, J., and Vassart, G.**, Somatic mutations in the thyrotropin receptor gene cause hyperfunctioning thyroid adenomas. *Nature*, 365(6447): 649–651. 1993.

[94] **Vasseur, C., Rodien, P., Beau, I., Desroches, A., Gérard, C., de Poncheville, L., Chaplot, S., Savagner, F., Croué, A., Mathieu, E., et al.**, A chorionic gonadotropin–sensitive mutation in the follicle-stimulating hormone receptor as a cause of familial gestational spontaneous ovarian hyperstimulation syndrome. *New England Journal of Medicine*, 349(8): 753–759. 2003.

[95] **Gregory, K.J., Dong, E.N., Meiler, J., and Conn, P.J.**, Allosteric modulation of metabotropic glutamate receptors: structural insights and therapeutic potential. *Neuropharmacology*, 60(1): 66–81. 2011.

[96] **Thathiah, A. and De Strooper, B.**, The role of G protein-coupled receptors in the pathology of Alzheimer's disease. *Nature Reviews Neuroscience*, 12(2): 73–87. 2011.

[97] **Dragun, D., Philippe, A., Catar, R., Hegner, B., et al.**, Autoimmune mediated G-protein receptor activation in cardiovascular and renal pathologies. *Thromb Haemost*, 101(4): 643–648. 2009.

[98] **Vischer, H., Vink, C., and Smit, M.**, A viral conspiracy: hijacking the chemokine system through virally encoded pirated chemokine receptors. In *Chemokines and Viral Infection*, pp. 121–154, Springer. 2006.

[99] **Lombardi, M.S., Kavelaars, A., and Heijnen, C.J.**, Role and modulation of G protein-coupled receptor signaling in inflammatory processes. *Critical Reviews in Immunology*, 22(2). 2002.

[100] **Penela, P., Murga, C., Ribas, C., Tutor, A.S., Peregrín, S., and Mayor, F.**, Mechanisms of regulation of G protein-coupled receptor kinases (GRKs) and cardiovascular disease. *Cardiovascular Research*, 69(1): 46–56. 2006.

[101] **Shichida, Y. and Imai, H.**, Visual pigment: G-protein-coupled receptor for light signals. *Cellular and Molecular Life Sciences CMLS*, 54(12): 1299–1315. 1998.

[102] **Lappano, R. and Maggiolini, M.**, G protein-coupled receptors: novel targets for drug discovery in cancer. *Nature Reviews Drug Discovery*, 10(1): 47–60. 2011.

[103] **Drews, J.**, Drug discovery: a historical perspective. *Science*, 287(5460): 1960–1964. 2000.

[104] **Overington, J.P., Al-Lazikani, B., and Hopkins, A.L.**, How many drug targets are there? *Nature Reviews Drug Discovery*, 5(12): 993–996. 2006.

[105] **Lin, S. and Civelli, O.**, Orphan G protein-coupled receptors: targets for new therapeutic interventions. *Annals of Medicine*, 36(3): 204–214. 2004.

[106] **Schiöth, H.B. and Fredriksson, R.**, The GRAFS classification system of G-protein coupled receptors in comparative perspective. *General and Comparative Endocrinology*, 142(1): 94–101. 2005.

[107] **Kolakowski Jr, L.F.**, GCRDb: a G-protein-coupled receptor database. *Receptors and Channels*, 2(1): 1–7. 1993.

[108] **Fredriksson, R., Lagerström, M.C., Lundin, L.G., and Schiöth, H.B.**, The G-protein-coupled receptors in the human genome form five main families. Phylogenetic analysis, paralogon groups, and fingerprints. *Molecular Pharmacology*, 63(6): 1256–1272. 2003.

[109] **Torphy, T.J.**, β-Adrenoceptors, cAMP and airway smooth muscle relaxation: challenges to the dogma. *Trends in Pharmacological Sciences*, 15(10): 370–374. 1994.

[110] **Follin-Arbelet, V., Hofgaard, P.O., Hauglin, H., Naderi, S., Sundan, A., Blomhoff, R., Bogen, B., and Blomhoff, H.K.**, Cyclic AMP induces apoptosis in multiple myeloma cells and inhibits tumor development in a mouse myeloma model. *BMC Cancer*, 11(1): 301. 2011.

[111] **Rocha, A.S., Paternot, S., Coulonval, K., Dumont, J.E., Soares, P., and Roger, P.P.**, Cyclic AMP inhibits the proliferation of thyroid carcinoma cell lines through regulation of CDK4 phosphorylation. *Molecular Biology of the Cell*, 19(11): 4814–4825. 2008.

[112] **Lounnas, V., Ritschel, T., Kelder, J., McGuire, R., Bywater, R.P., and Foloppe, N.**, Current progress in structure-based rational drug design marks a new mindset in drug discovery. *Computational and Structural Biotechnology Journal*, 5(6): 1–14. 2013.

[113] **Smith, S.O.**, Structure and activation of the visual pigment rhodopsin. *Annual Review of Biophysics*, 39: 309–328. 2010.

[114] **Provasi, D., Artacho, M.C., Negri, A., Mobarec, J.C., and Filizola, M.**, Ligand-induced modulation of the free-energy landscape of G protein-coupled receptors explored by adaptive biasing techniques. *PLoS Comput Biol*, 7(10): e1002193. 2011.

[115] **Ye, S., Zaitseva, E., Caltabiano, G., Schertler, G.F., Sakmar, T.P., Deupi, X., and Vogel, R.**, Tracking G-protein-coupled receptor activation using genetically encoded infrared probes. *Nature*, 464(7293): 1386–1389. 2010.

[116] **Deupi, X. and Kobilka, B.**, Activation of G protein–coupled receptors. *Advances in Protein Chemistry*, 74: 137–166. 2007.

[117] **Ahuja, S. and Smith, S.O.**, Multiple switches in G protein-coupled receptor activation. *Trends in Pharmacological Sciences*, 30(9): 494–502. 2009.

[118] **Hofmann, K.P., Scheerer, P., Hildebrand, P.W., Choe, H.W., Park, J.H., Heck, M., and Ernst, O.P.**, A G protein-coupled receptor at work: the rhodopsin model. *Trends in Biochemical Sciences*, 34(11): 540–552. 2009.

[119] **Oliveira, L., Paiva, P.B., Paiva, A., and Vriend, G.**, Identification of functionally conserved residues with the use of entropy–variability plots. *Proteins: Structure, Function, and Bioinformatics*, 52(4): 544–552. 2003.

[120] **Rubenstein, L.A. and Lanzara, R.G.**, Activation of G protein-coupled receptors entails cysteine modulation of agonist binding. *Journal of Molecular Structure: THEOCHEM*, 430: 57–71. 1998.

[121] **Sanders, M.P., Fleuren, W.W., Verhoeven, S., van den Beld, S., Alkema, W., de Vlieg, J., and Klomp, J.P.**, ss-TEA: Entropy based identification of receptor specific ligand binding residues from a multiple sequence alignment of class A GPCRs. *BMC Bioinformatics*, 12(1): 1. 2011.

[122] **Provasi, D., Artacho, M.C., Negri, A., Mobarec, J.C., and Filizola, M.**, Ligand-induced modulation of the free-energy landscape of G protein-coupled receptors explored by adaptive biasing techniques. *PLoS Computational Biology*, 7(10): e1002193. 2011.

[123] **Jansen, A.S., Van Nguyen, X., Karpitskiy, V., Mettenleiter, T.C., and Loewy, A.D.**, Central command neurons of the sympathetic nervous system: basis of the fight-or-flight response. *Science*, 270(5236): 644–646. 1995.

[124] **Xiao, R.P., Avdonin, P., Zhou, Y.Y., Cheng, H., Akhter, S.A., Eschenhagen, T., Lefkowitz, R.J., Koch, W.J., and Lakatta, E.G.**, Coupling of β2-adrenoceptor to Gi proteins and its physiological relevance in murine cardiac myocytes. *Circulation Research*, 84(1): 43–52. 1999.

[125] **Azzi, M., Charest, P.G., Angers, S., Rousseau, G., Kohout, T., Bouvier, M., and Piñeyro, G.**, β-Arrestin-mediated activation of MAPK by inverse agonists reveals distinct active conformations for G protein-coupled receptors. *Proceedings of the National Academy of Sciences*, 100(20): 11406–11411. 2003.

[126] **Shenoy, S.K., Drake, M.T., Nelson, C.D., Houtz, D.A., Xiao, K., Madabushi, S., Reiter, E., Premont, R.T., Lichtarge, O., and Lefkowitz, R.J.**, β-Arrestin-dependent, G protein-independent ERK1/2 activation by the β2 adrenergic receptor. *Journal of Biological Chemistry*, 281(2): 1261–1273. 2006.

[127] **Kobilka, B.**, The structural basis of G-protein-coupled receptor signaling (Nobel Lecture). *Angewandte Chemie International Edition*, 52(25): 6380–6388. 2013.

[128] **Manglik, A. and Kobilka, B.**, The role of protein dynamics in GPCR function: Insights from the β2AR and rhodopsin. *Current Opinion in Cell Biology*, 27: 136–143. 2014.

[129] **Dawaliby, R., Trubbia, C., Delporte, C., Masureel, M., Van Antwerpen, P., Kobilka, B.K., and Govaerts, C.**, Allosteric regulation of G protein-coupled receptor activity by phospholipids. *Nature Chemical Biology*, 12(1): 35–39. 2016.

[130] **Azpiazu, I. and Gautam, N.**, A fluorescence resonance energy transfer-based sensor indicates that receptor access to a G protein is unrestricted in a living mammalian cell. *Journal of Biological Chemistry*, 279(26): 27709–27718. 2004.

[131] **Levitzki, A.**, From epinephrine to cyclic AMP. *Science*, 241(4867): 800–806. 1988.

[132] **Ostrom, R.S. and Insel, P.A.**, The evolving role of lipid rafts and caveolae in G protein-coupled receptor signaling: Implications for molecular pharmacology. *British Journal of Pharmacology*, 143(2): 235–245. 2004.

[133] **Insel, P.A., Head, B.P., Ostrom, R.S., Patel, H.H., Swaney, J.S., Tang, C.M., and Roth, D.M.**, Caveolae and lipid rafts: G protein-coupled receptor signaling microdomains in cardiac myocytes. *Annals of the New York Academy of Sciences*, 1047(1): 166–172. 2005.

[134] **Couet, J., Li, S., Okamoto, T., Ikezu, T., and Lisanti, M.P.**, Identification of peptide and protein ligands for the caveolin-scaffolding domain. Implications for the interaction of caveolin with caveolae-associated proteins. *Journal of Biological Chemistry*, 272(10): 6525–6533. 1997.

[135] **Murthy, K., Teng, B.Q., Zhou, H., Jin, J.G., Grider, J., and Makhlouf, G.**, Gi-1/Gi-2-dependent signaling by single-transmembrane natriuretic peptide clearance receptor. *American Journal of Physiology–Gastrointestinal and Liver Physiology*, 278(6): G974–G980. 2000.

[136] **Huber, T., Botelho, A.V., Beyer, K., and Brown, M.F.**, Membrane model for the G-protein-coupled receptor rhodopsin: hydrophobic interface and dynamical structure. *Biophysical Journal*, 86(4): 2078–2100. 2004.

[137] **Barnett-Norris, J., Lynch, D., and Reggio, P.H.**, Lipids, lipid rafts and caveolae: their importance for GPCR signaling and their centrality to the endocannabinoid system. *Life Sciences*, 77(14): 1625–1639. 2005.

[138] **Chini, B. and Parenti, M.**, G-protein coupled receptors in lipid rafts and caveolae: how, when and why do they go there? *Journal of Molecular Endocrinology*, 32(2): 325–338. 2004.

[139] **Chini, B. and Parenti, M.**, G-protein-coupled receptors, cholesterol and palmitoylation: facts about fats. *Journal of Molecular Endocrinology*, 42(5): 371–379. 2009.

[140] **Renner, U., Glebov, K., Lang, T., Papusheva, E., Balakrishnan, S., Keller, B., Richter, D.W., Jahn, R., and Ponimaskin, E.**, Localization of the mouse 5-hydroxytryptamine 1A receptor in lipid microdomains depends on its palmitoylation and is involved in receptor-mediated signaling. *Molecular Pharmacology*, 72(3): 502–513. 2007.

[141] **Paila, Y.D. and Chattopadhyay, A.**, Membrane cholesterol in the function and organization of G-protein coupled receptors. In *Cholesterol Binding and Cholesterol Transport Proteins:*, pp. 439–466, Springer. 2010.

[142] **Gimpl, G., Wiegand, V., Burger, K., and Fahrenholz, F.**, Cholesterol and steroid hormones: modulators of oxytocin receptor function. *Progress in Brain Research*, 139: 43–55. 2002.

[143] **Gimpl, G., Burger, K., and Fahrenholz, F.**, A closer look at the cholesterol sensor. *Trends in Biochemical Sciences*, 27(12): 596–599. 2002.

[144] **Ohvo-Rekilä, H., Ramstedt, B., Leppimäki, P., and Slotte, J.P.**, Cholesterol interactions with phospholipids in membranes. *Progress in Lipid Research*, 41(1): 66–97. 2002.

[145] **Risselada, H.J. and Marrink, S.J.**, The molecular face of lipid rafts in model membranes. *Proceedings of the National Academy of Sciences*, 105(45): 17367–17372. 2008.

[146] **Schäfer, L.V., de Jong, D.H., Holt, A., Rzepiela, A.J., de Vries, A.H., Poolman, B., Killian, J.A., and Marrink, S.J.**, Lipid packing drives the segregation of transmembrane helices into disordered lipid domains in model membranes. *Proceedings of the National Academy of Sciences*, 108(4): 1343–1348. 2011.

[147] **de Jong, D.H., Lopez, C.A., and Marrink, S.J.**, Molecular view on protein sorting into liquid-ordered membrane domains mediated by gangliosides and lipid anchors. *Faraday Discussions*, 161: 347–363. 2013.

[148] **Seno, K., Kishimoto, M., Abe, M., Higuchi, Y., Mieda, M., Owada, Y., Yoshiyama, W., Liu, H., and Hayashi, F.**, Light- and guanosine 5'-3-0-(thio) triphosphate-sensitive localization of a G protein and its effector on detergent-resistant membrane rafts in rod photoreceptor outer segments. *Journal of Biological Chemistry*, 276(24): 20813–20816. 2001.

[149] **Karplus, M. and McCammon, J.A.**, Molecular dynamics simulations of biomolecules. *Nature Structural and Molecular Biology*, 9(9): 646–652. 2002.

[150] **Nair, P.C. and Miners, J.O.**, Molecular dynamics simulations: from structure function relationships to drug discovery. *In Silico Pharmacology*, 2(1): 1. 2014.

[151] **Marrink, S.J., de Vries, A.H., and Tieleman, D.P.**, Lipids on the move: simulations of membrane pores, domains, stalks and curves. *Biochimica et Biophysica Acta (BBA)-Biomembranes*, 1788(1): 149–168. 2009.

[152] **Hansson, T., Oostenbrink, C., and van Gunsteren, W.**, Molecular dynamics simulations. *Current Opinion in Structural Biology*, 12(2): 190–196. 2002.

[153] **Abraham, M., Van de Spoel, D., Lindahl, E., Hess, B., and the GROMACS development team**, GROMACS user manual version 5.0.7. Accessed 14-April-2016.

[154] **Orozco, M., Orellana, L., Naganathan, A., Emperador, A., Carrillo, O., Gelpí, J., et al.**, Coarse-grained representation of protein flexibility. Foundations, successes, and shortcomings. *Advances in Protein Chemistry and Structural Biology*, 85: 183–215. 2011.

[155] **Van Duin, A.C., Dasgupta, S., Lorant, F., and Goddard, W.A.**, ReaxFF: a reactive force field for hydrocarbons. *The Journal of Physical Chemistry A*, 105(41): 9396–9409. 2001.

[156] **Lopes, P.E., Huang, J., Shim, J., Luo, Y., Li, H., Roux, B., and MacKerell Jr, A.D.**, Polarizable force field for peptides and proteins based on the classical drude oscillator. *Journal of Chemical Theory and Computation*, 9(12): 5430–5449. 2013.

[157] **Kaminski, G.A., Friesner, R.A., Tirado-Rives, J., and Jorgensen, W.L.**, Evaluation and reparametrization of the OPLS-AA force field for proteins via comparison with accurate quantum chemical calculations on peptides. *The Journal of Physical Chemistry B*, 105(28): 6474–6487. 2001.

[158] **Marrink, S.J., De Vries, A.H., and Mark, A.E.**, Coarse grained model for semiquantitative lipid simulations. *The Journal of Physical Chemistry B*, 108(2): 750–760. 2004.

[159] **Berendsen, H.J.**, *Simulating the Physical World: Hierarchical Modeling from Quantum Mechanics to Fluid Dynamics*. Cambridge University Press. 2007, 139–210 pp.

[160] **Toukmaji, A.Y. and Board, J.A.**, Ewald summation techniques in perspective: a survey. *Computer Physics Communications*, 95(2): 73–92. 1996.

[161] **Van Gunsteren, W. and Berendsen, H.**, A leap-frog algorithm for stochastic dynamics. *Molecular Simulation*, 1(3): 173–185. 1988.

[162] **Lane, T.J., Shukla, D., Beauchamp, K.A., and Pande, V.S.**, To milliseconds and beyond: challenges in the simulation of protein folding. *Current opinion in structural biology*, 23(1): 58–65. 2013.

[163] **Marrink, S.J., Risselada, H.J., Yefimov, S., Tieleman, D.P., and De Vries, A.H.**, The MARTINI force field: coarse grained model for biomolecular simulations. *The Journal of Physical Chemistry B*, 111(27): 7812–7824. 2007.

[164] **Marrink, S.J. and Tieleman, D.P.**, Perspective on the Martini model. *Chemical Society Reviews*, 42(16): 6801–6822. 2013.

[165] **Monticelli, L., Kandasamy, S.K., Periole, X., Larson, R.G., Tieleman, D.P., and Marrink, S.J.**, The MARTINI coarse-grained force field: extension to proteins. *Journal of Chemical Theory and Computation*, 4(5): 819–834. 2008.

[166] **Periole, X., Cavalli, M., Marrink, S.J., and Ceruso, M.A.**, Combining an elastic network with a coarse-grained molecular force field: structure, dynamics, and intermolecular recognition. *Journal of Chemical Theory and Computation*, 5(9): 2531–2543. 2009.

[167] **Wassenaar, T.A., Pluhackova, K., Böckmann, R.A., Marrink, S.J., and Tieleman, D.P.**, Going backward: a flexible geometric approach to reverse transformation from coarse grained to atomistic models. *Journal of Chemical Theory and Computation*, 10(2): 676–690. 2014.

[168] **De Jong, T., Linn, M.C., and Zacharia, Z.C.**, Physical and virtual laboratories in science and engineering education. *Science*, 340(6130): 305–308. 2013.

[169] **Jo, S., Kim, T., Iyer, V.G., and Im, W.**, CHARMM-GUI: a web-based graphical user interface for CHARMM. *Journal of Computational Chemistry*, 29(11): 1859–1865. 2008.

[170] **Qi, Y., Ingólfsson, H.I., Cheng, X., Lee, J., Marrink, S.J., and Im, W.**, CHARMM-GUI Martini maker for coarse-grained simulations with the Martini force field. *Journal of Chemical Theory and Computation*, 11(9): 4486–4494. 2015.

[171] **Daily, M.D., Olsen, B.N., Schlesinger, P.H., Ory, D.S., and Baker, N.A.**, Improved coarse-grained modeling of cholesterol-containing lipid bilayers. *Journal of Chemical Theory and Computation*, 10(5): 2137–2150. 2014.

[172] **Melo, M., Ingólfsson, H., and Marrink, S.**, Parameters for Martini sterols and hopanoids based on a virtual-site description. *The Journal of Chemical Physics*, 143(24): 243152. 2015.

[173] **Kulig, W., Olżyńska, A., Jurkiewicz, P., Kantola, A.M., Komulainen, S., Manna, M., Pourmousa, M., Vazdar, M., Cwiklik, L., Rog, T., et al.**, Cholesterol under oxidative stress – How lipid membranes sense oxidation as cholesterol is being replaced by oxysterols. *Free Radical Biology and Medicine*, 84: 30–41. 2015.

[174] **Manna, M., Kulig, W., Javanainen, M., Tynkkynen, J., Hensen, U., Müller, D.J., Rog, T., and Vattulainen, I.**, How to minimize artifacts in atomistic simulations of membrane proteins, whose crystal structure is heavily engineered: β2-adrenergic receptor in the spotlight. *Journal of Chemical Theory and Computation*, 11(7): 3432–3445. 2015.

[175] **Kulig, W., Tynkkynen, J., Javanainen, M., Manna, M., Rog, T., Vattulainen, I., and Jungwirth, P.**, How well does cholesteryl hemisuccinate mimic cholesterol in saturated phospholipid bilayers? *Journal of Molecular Modeling*, 20(2): 1–9. 2014.

[176] **Zhang, Y., Carter, J.W., Lervik, A., Brooks, N.J., Seddon, J.M., and Bresme, F.**, Structural organization of sterol molecules in DPPC bilayers: a coarse-grained molecular dynamics investigation. *Soft Matter*. 2016.

[177] **Baoukina, S. and Tieleman, D.P.**, Computer simulations of phase separation in lipid bilayers and monolayers. *Methods in Membrane Lipids*, pp. 307–322. 2015.

[178] **Waheed, Q., Tjörnhammar, R., and Edholm, O.**, Phase transitions in coarse-grained lipid bilayers containing cholesterol by molecular dynamics simulations. *Biophysical Journal*, 103(10): 2125–2133. 2012.

[179] **Wassenaar, T.A., Ingólfsson, H.I., Böckmann, R.A., Tieleman, D.P., and Marrink, S.J.**, Computational lipidomics with insane: a versatile tool for generating custom membranes for molecular simulations. *Journal of Chemical Theory and Computation*, 11(5): 2144–2155. 2015.

[180] **Maciejewski, A., Pasenkiewicz-Gierula, M., Cramariuc, O., Vattulainen, I., and Rog, T.**, Refined OPLS all-atom force field for saturated phosphatidylcholine bilayers at full hydration. *The Journal of Physical Chemistry B*, 118(17): 4571–4581. 2014.

[181] **Kulig, W., Pasenkiewicz-Gierula, M., and Róg, T.**, Cis and trans unsaturated phosphatidylcholine bilayers: a molecular dynamics simulation study. *Chemistry and Physics of Lipids*, 195: 12–20. 2016.

[182] **Kulig, W., Pasenkiewicz-Gierula, M., and Róg, T.**, Topologies, structures and parameter files for lipid simulations in GROMACS with the OPLS-aa force field: DPPC, POPC, DOPC, PEPC, and cholesterol. *Data in Brief*, 5: 333–336. 2015.

[183] **Jorgensen, W.L., Chandrasekhar, J., Madura, J.D., Impey, R.W., and Klein, M.L.**, Comparison of simple potential functions for simulating liquid water. *The Journal of Chemical Physics*, 79(2): 926–935. 1983.

[184] **Abraham, M.J., Murtola, T., Schulz, R., Páll, S., Smith, J.C., Hess, B., and Lindahl, E.**, GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX*, 1: 19–25. 2015.

[185] **Berendsen, H.J., Postma, J.v., van Gunsteren, W.F., DiNola, A., and Haak, J.**, Molecular dynamics with coupling to an external bath. *The Journal of Chemical Physics*, 81(8): 3684–3690. 1984.

[186] **Bussi, G., Donadio, D., and Parrinello, M.**, Canonical sampling through velocity rescaling. *The Journal of Chemical Physics*, 126(1): 014101. 2007.

[187] **Parrinello, M. and Rahman, A.**, Polymorphic transitions in single crystals: A new molecular dynamics method. *Journal of Applied Physics*, 52(12): 7182–7190. 1981.

[188] **Essmann, U., Perera, L., Berkowitz, M.L., Darden, T., Lee, H., and Pedersen, L.G.**, A smooth particle mesh Ewald method. *The Journal of Chemical Physics*, 103(19): 8577–8593. 1995.

[189] **Hess, B., Bekker, H., Berendsen, H.J., Fraaije, J.G., et al.**, LINCS: a linear constraint solver for molecular simulations. *Journal of Computational Chemistry*, 18(12): 1463–1472. 1997.

[190] **Humphrey, W., Dalke, A., and Schulten, K.**, VMD: visual molecular dynamics. *Journal of Molecular Graphics*, 14(1): 33–38. 1996.

[191] **Gapsys, V., de Groot, B.L., and Briones, R.**, Computational analysis of local membrane properties. *Journal of Computer-aided Molecular Design*, 27(10): 845–858. 2013.

[192] **Huang, J., Buboltz, J.T., and Feigenson, G.W.**, Maximum solubility of cholesterol in phosphatidylcholine and phosphatidylethanolamine bilayers. *Biochimica et Biophysica Acta (BBA)-Biomembranes*, 1417(1): 89–100. 1999.

[193] **Gater, D.L., Saurel, O., Iordanov, I., Liu, W., Cherezov, V., and Milon, A.**, Two classes of cholesterol binding sites for the β2AR revealed by thermostability and NMR. *Biophysical Journal*, 107(10): 2305–2312. 2014.

[194] **Feller, S.E., Gawrisch, K., and Woolf, T.B.**, Rhodopsin exhibits a preference for solvation by polyunsaturated docosohexaenoic acid. *Journal of the American Chemical Society*, 125(15): 4434–4435. 2003.

[195] **Pitman, M.C., Grossfield, A., Suits, F., and Feller, S.E.**, Role of cholesterol and polyunsaturated chains in lipid-protein interactions: molecular dynamics simulation of rhodopsin in a realistic membrane environment. *Journal of the American Chemical Society*, 127(13): 4576–4577. 2005.

[196] **Grossfield, A., Feller, S.E., and Pitman, M.C.**, Contribution of omega-3 fatty acids to the thermodynamics of membrane protein solvation. *The Journal of Physical Chemistry B*, 110(18): 8907–8909. 2006.

[197] **Vidal, A. and McIntosh, T.J.**, Transbilayer peptide sorting between raft and nonraft bilayers: comparisons of detergent extraction and confocal microscopy. *Biophysical Journal*, 89(2): 1102–1108. 2005.

APPENDIX A. CONCENTRATION OF LIPIDS IN THE PHASE SEGREGATED MEMBRANE

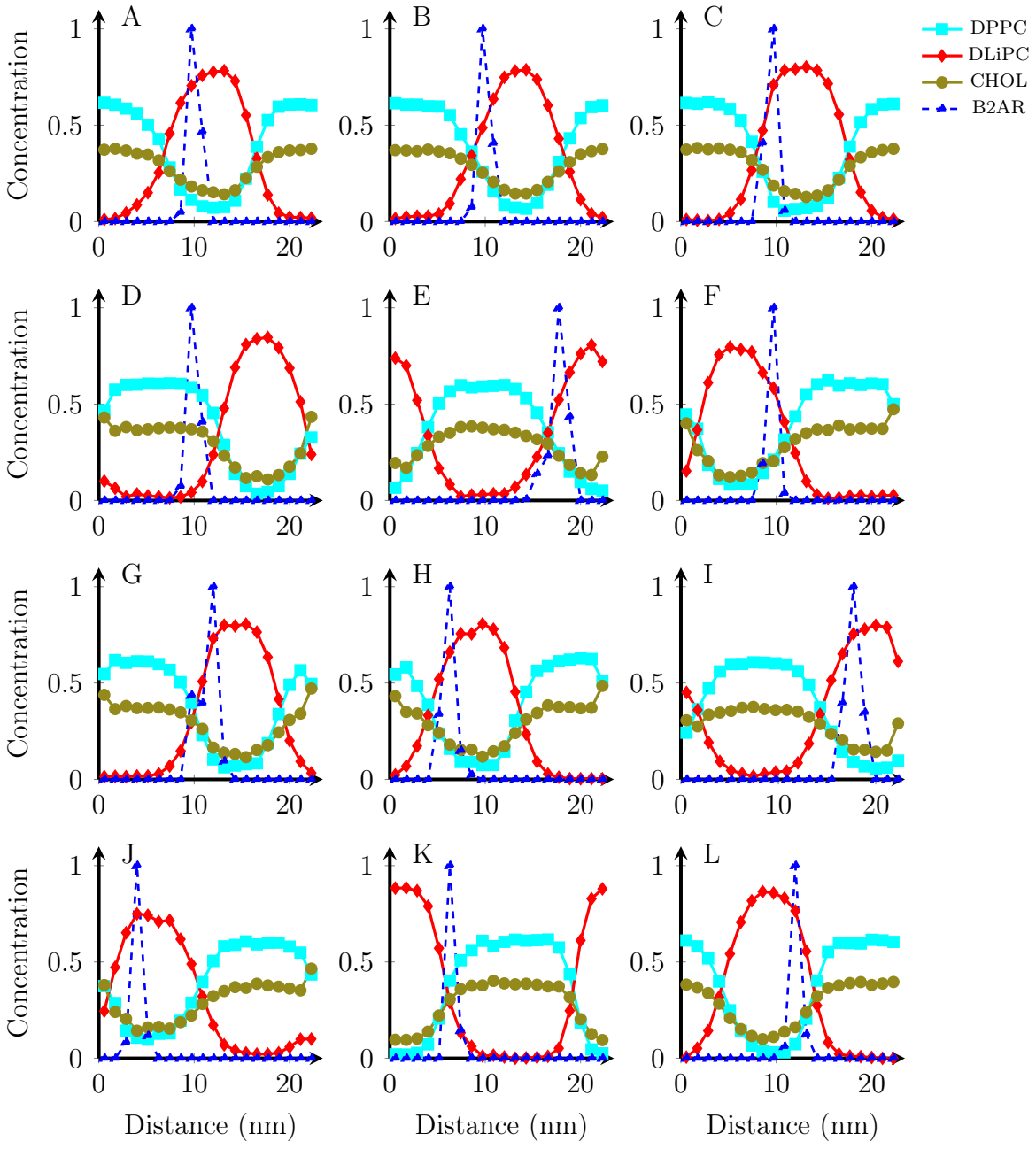


Figure A.1: The lipid concentration along the axis perpendicular to the lipid phase boundary in the systems (A-J) R1-R10 and (K-L) RC1-RC2.

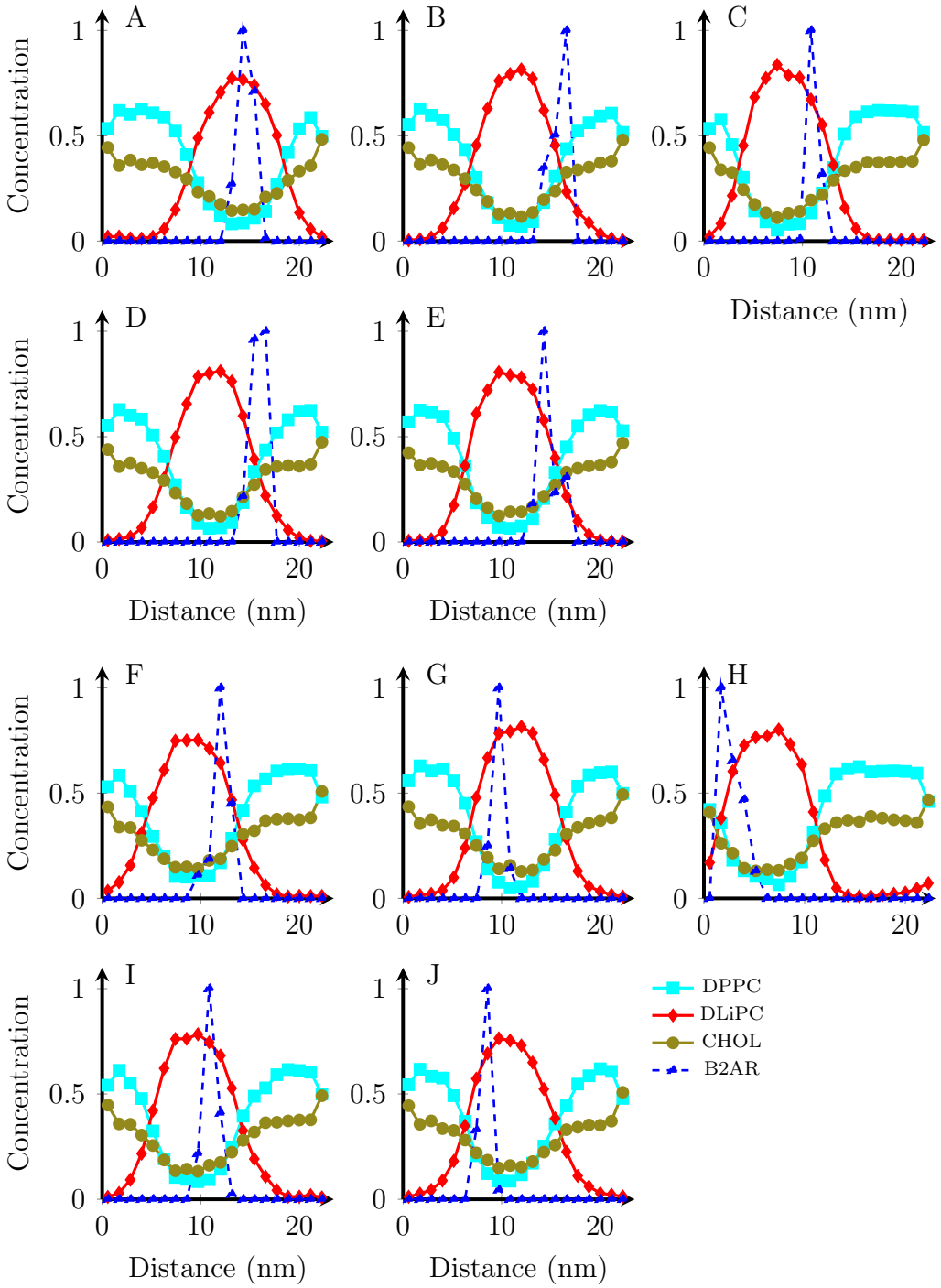


Figure A.2: The lipid concentration along the axis perpendicular to the lipid phase boundary in the systems (A-E) O1-O5 and (F-J) D1-D5.

APPENDIX B. CONCENTRATION OF LIPIDS AT INCREASING DISTANCES FROM THE PROTEIN

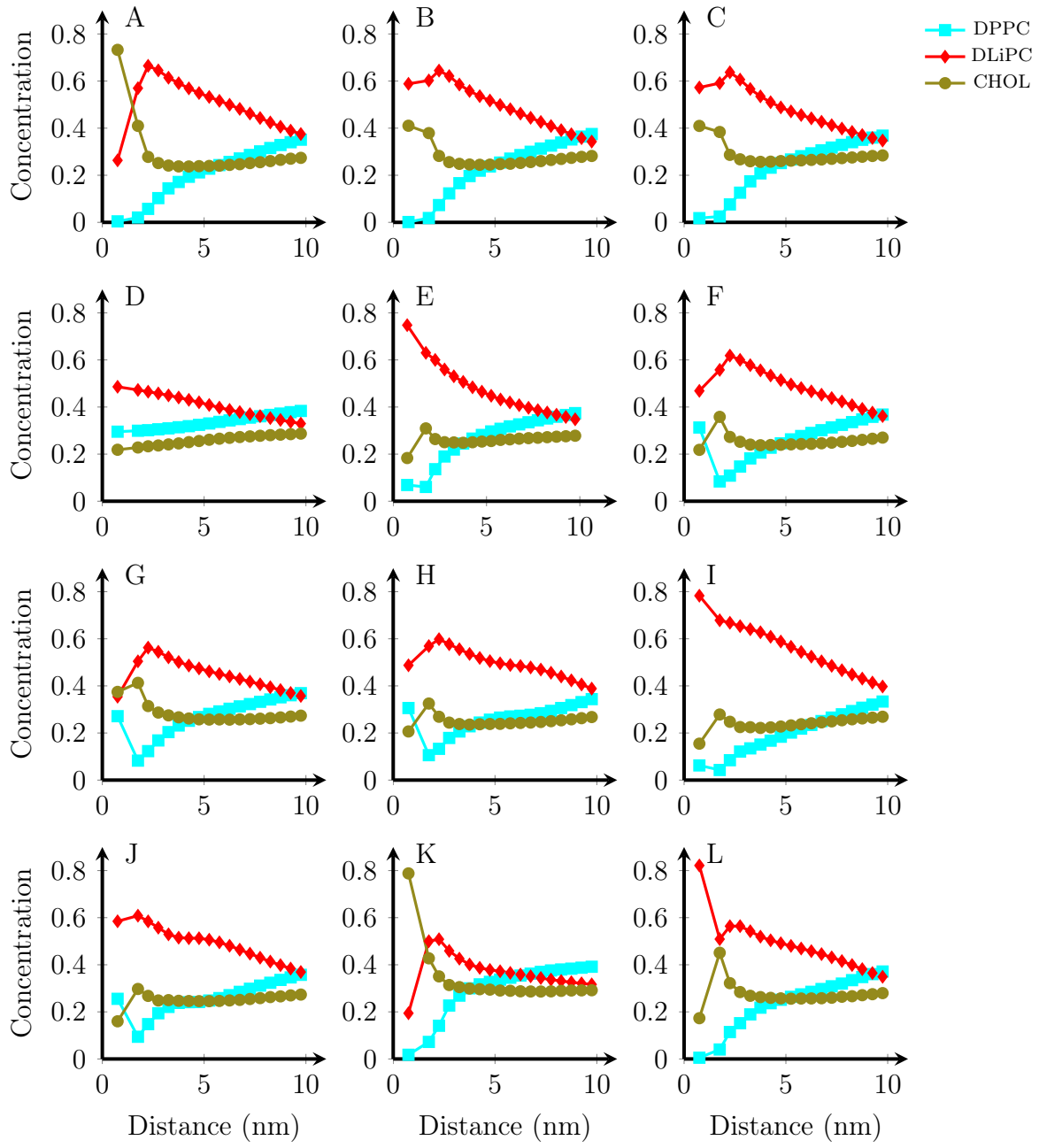


Figure B.1: The lipid concentration at distances [0..1.5,1.5..2.0,2.0..2.5,...,9.5..10] from protein in the systems (A-J) R1-R10 and (K-L) RC1-RC2.

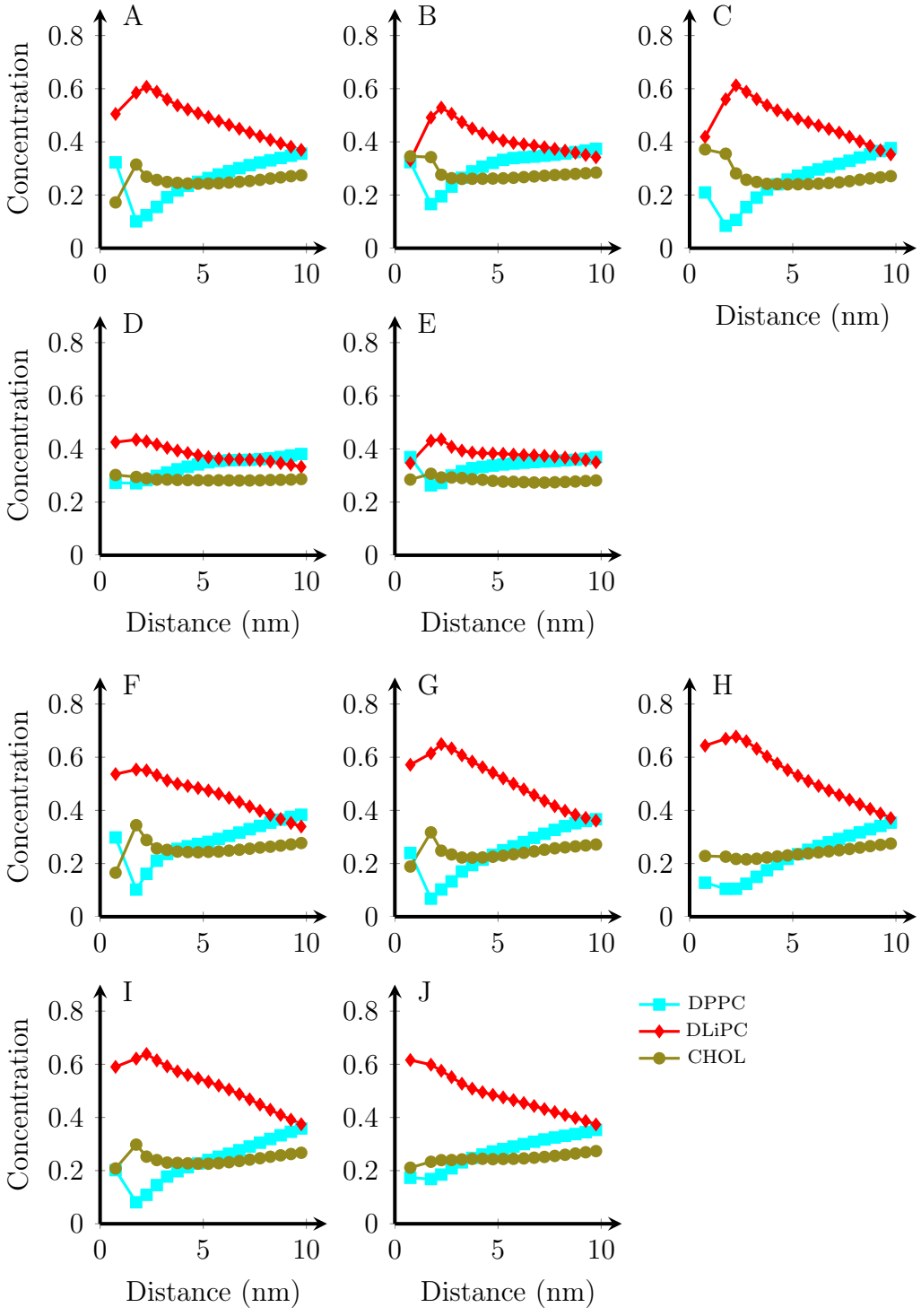


Figure B.2: The lipid concentration at distances [0..1.5,1.5..2.0,2.0..2.5,...,9.5..10] from protein in the systems (A-E) O1-O5 and (F-J) D1-D5.

APPENDIX C. RADIAL DISTRIBUTION OF LIPIDS FROM THE PROTEIN

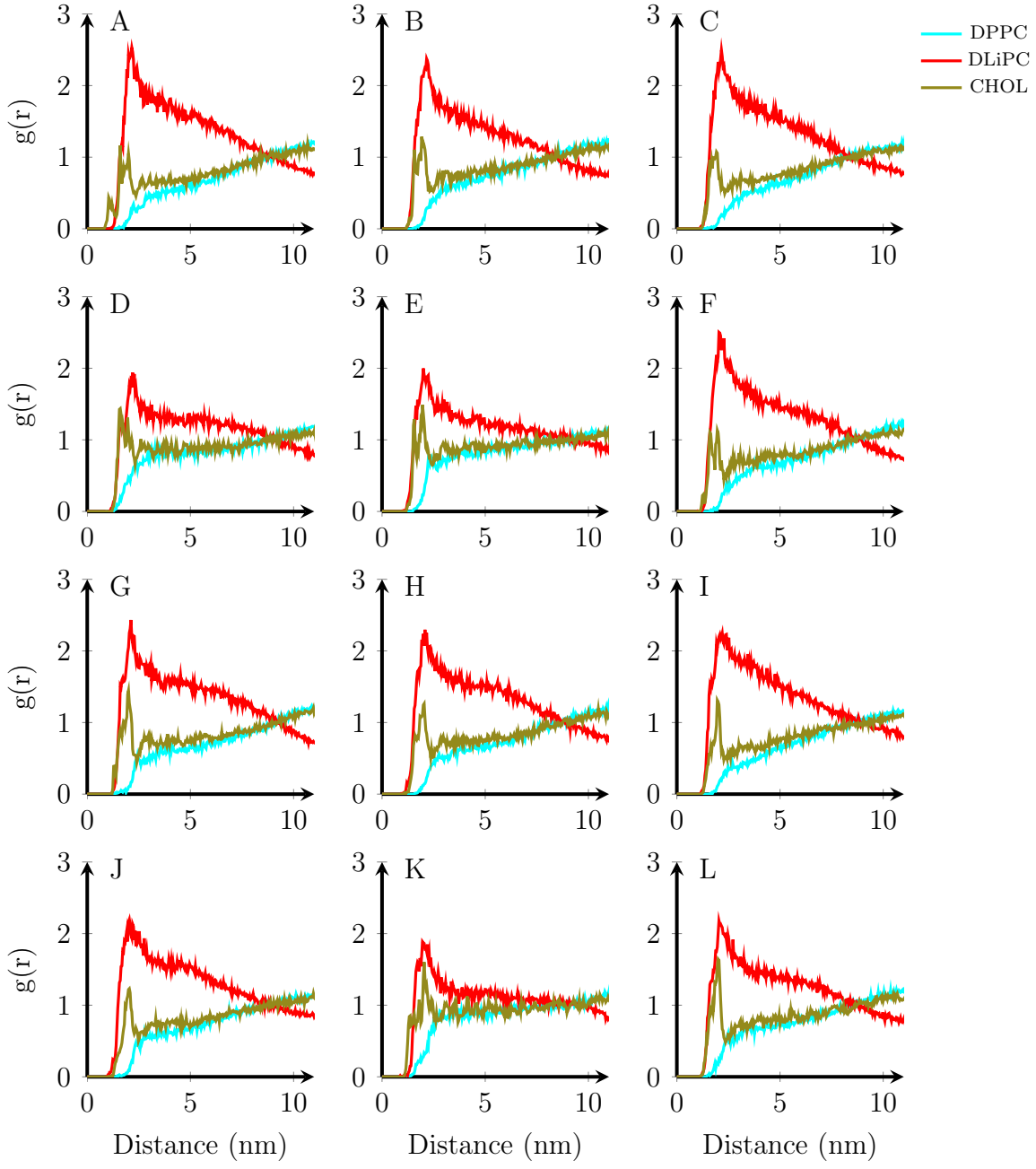


Figure C.1: The final RDF of lipids in the systems (A-J) R1-R10 and (K-L) RC1-RC2.

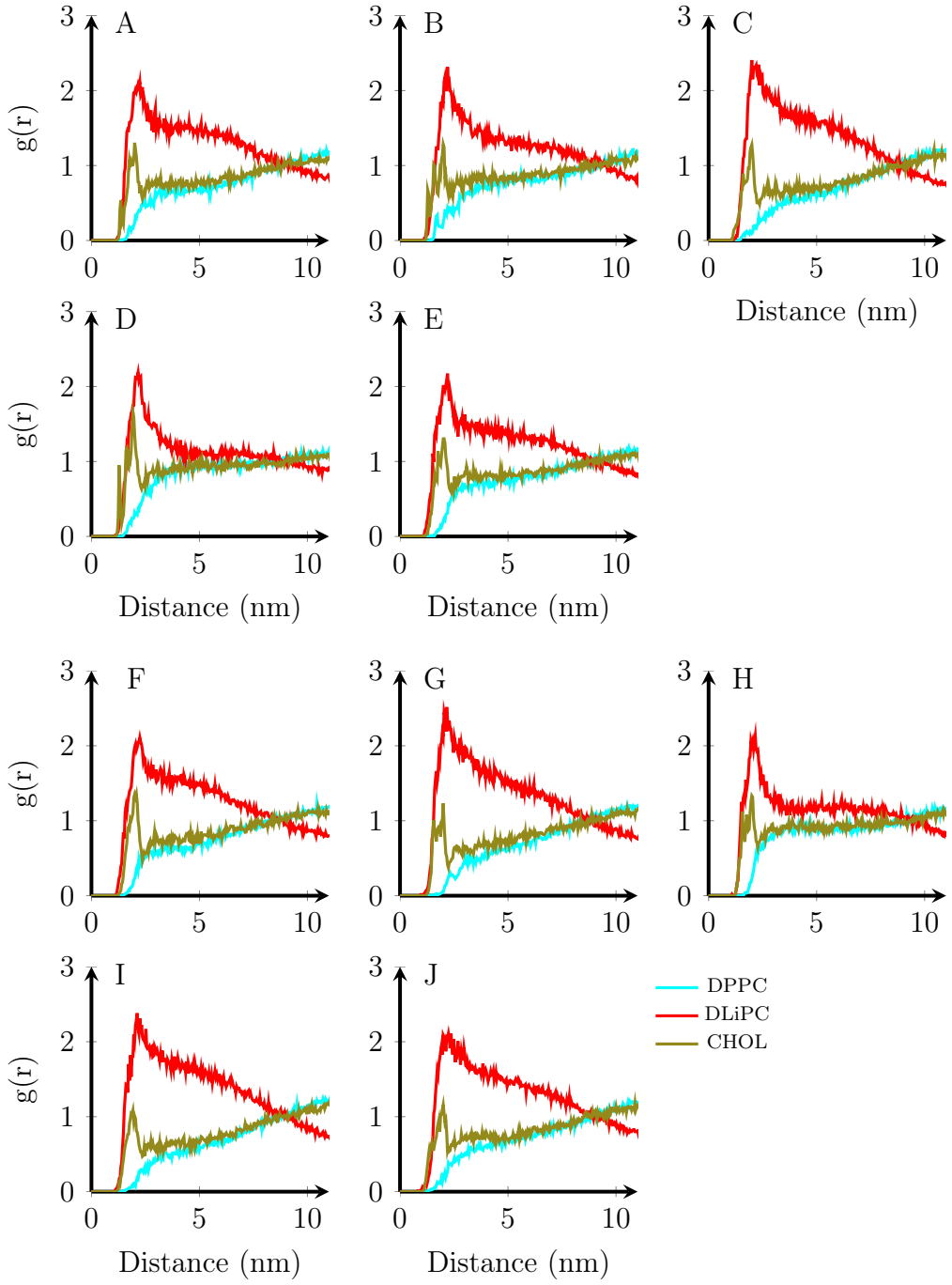


Figure C.2: The final RDF of lipids in the systems (A-E) O1-O5 and (F-J) D1-D5.

APPENDIX D. DIFFUSION COEFFICIENTS OF LIPIDS

Table D.1: The diffusion coefficients of the PC lipids fitted to the MSD curve calculated from the last 5 μs trajectory of each system.

	DLiPC		**DPPC**	
System	D (1e-8 cm^2/s)	Error	D (1e-8 cm^2/s)	Error
R1	4.24	0.01	1.81	0.01
R2	4.44	0.03	1.76	0.01
R3	4.63	0.01	1.52	0.01
R4	4.51	0.02	1.79	0.01
R5	5.46	0.02	1.68	0.01
R6	4.49	0.01	1.79	0.01
R7	4.56	0.01	1.88	0.01
R8	4.63	0.01	1.59	0.01
R9	3.28	0.01	1.76	0.01
R10	3.94	0.02	1.63	0.01
Mean & av.dev.	**4.42**	**0.55**	**1.72**	**0.12**
RC1	6.83	0.02	0.78	0.01
RC2	5.99	0.01	0.79	0.01
Mean & av.dev.	**6.41**	**0.59**	**0.79**	**0.01**
O1	5.28	0.02	1.50	0.01
O2	5.12	0.01	1.63	0.01
O3	4.53	0.01	1.63	0.01
O4	5.78	0.01	1.68	0.01
O5	5.03	0.01	1.72	0.01
Mean & av.dev.	**5.15**	**0.45**	**1.63**	**0.09**
D1	4.98	0.02	1.55	0.01
D2	4.86	0.02	1.67	0.01
D3	4.85	0.02	1.75	0.01
D4	5.63	0.02	1.79	0.01
D5	3.86	0.02	1.56	0.01
Mean & av.dev.	**4.84**	**0.64**	**1.66**	**0.11**

APPENDIX E. LIPID DENSITY MAPS AND NUMBER OF CONTACTS IN SYSTEMS RC1–RC2

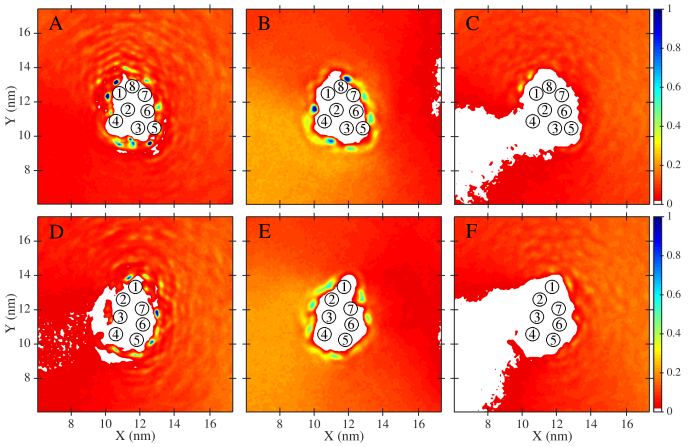


Figure E.1: The number density of (A) CHOL, (B) DLiPC, and (C) DPPC on the intracellular leaflet of the bilayer and (D-E) on the extracellular leaflet of the bilayer respectively, averaged over of systems RC1-RC2.

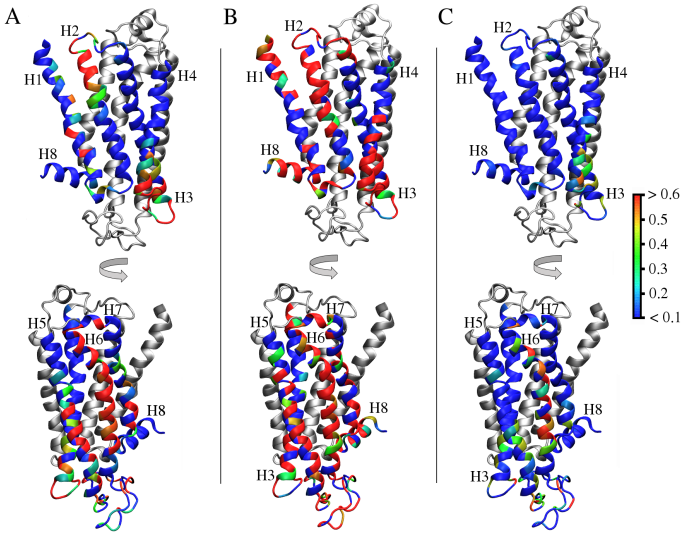


Figure E.2: The lipid occupancy time of (A) CHOL, (B) DLiPC, and (C) DPPC on the β_2AR surface averaged over systems RC1-RC2.

APPENDIX F. TWO-DIMENSIONAL THICKNESS AND LIPID CHAIN ORDER PARAMETER

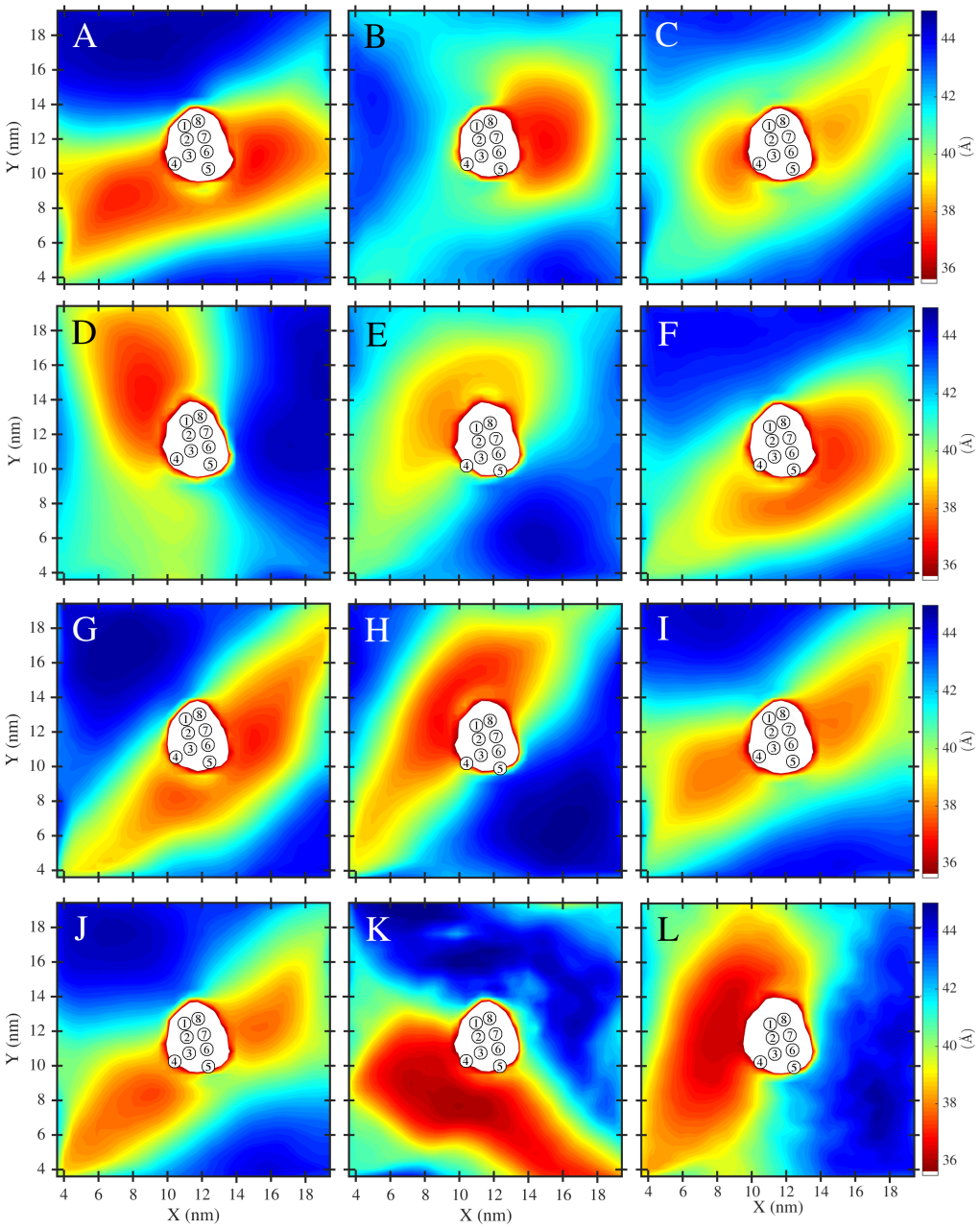


Figure F.1: The two-dimensional bilayer thickness in systems (A-J) R1-R10 and (K-L) RC1-RC2.

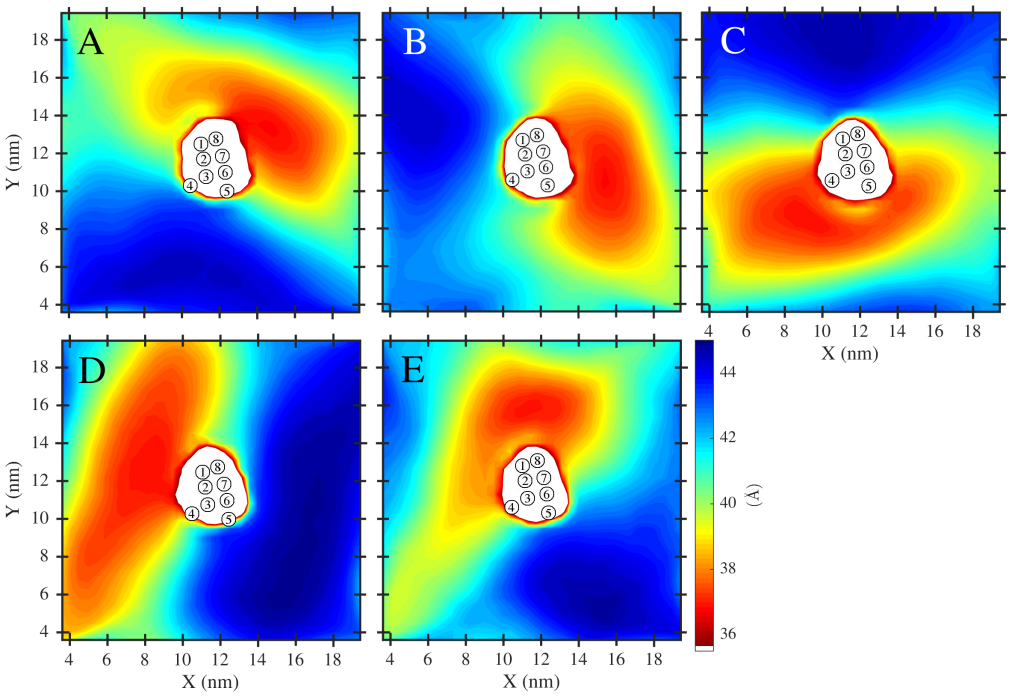


Figure F.2: The two-dimensional bilayer thickness in systems (A-E) O1-O5.

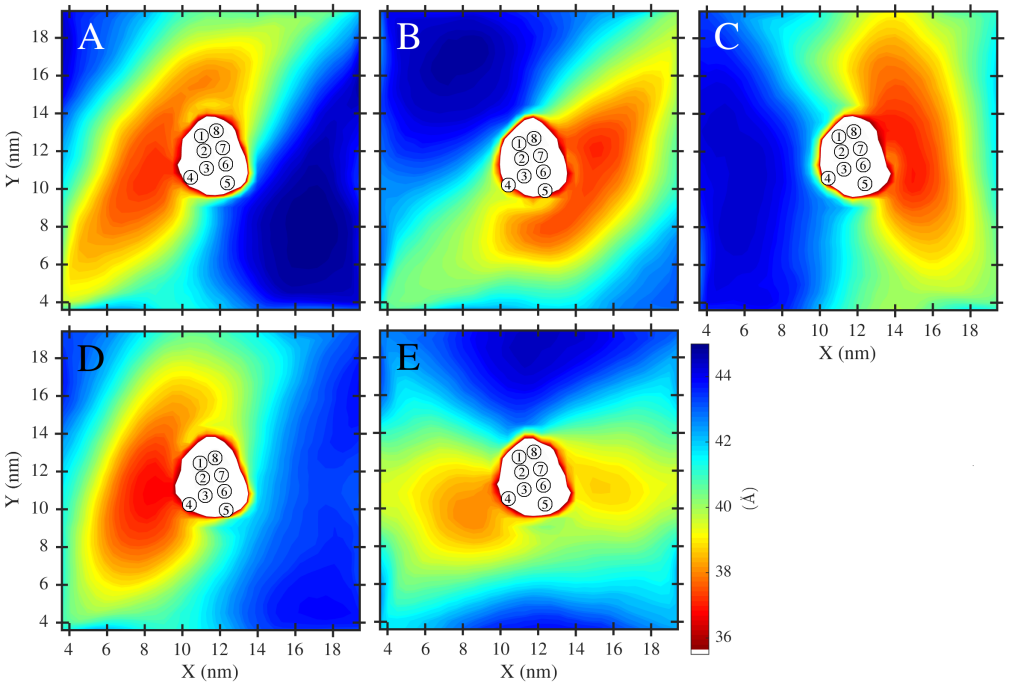


Figure F.3: The two-dimensional bilayer thickness in systems (A-E) D1-D5.

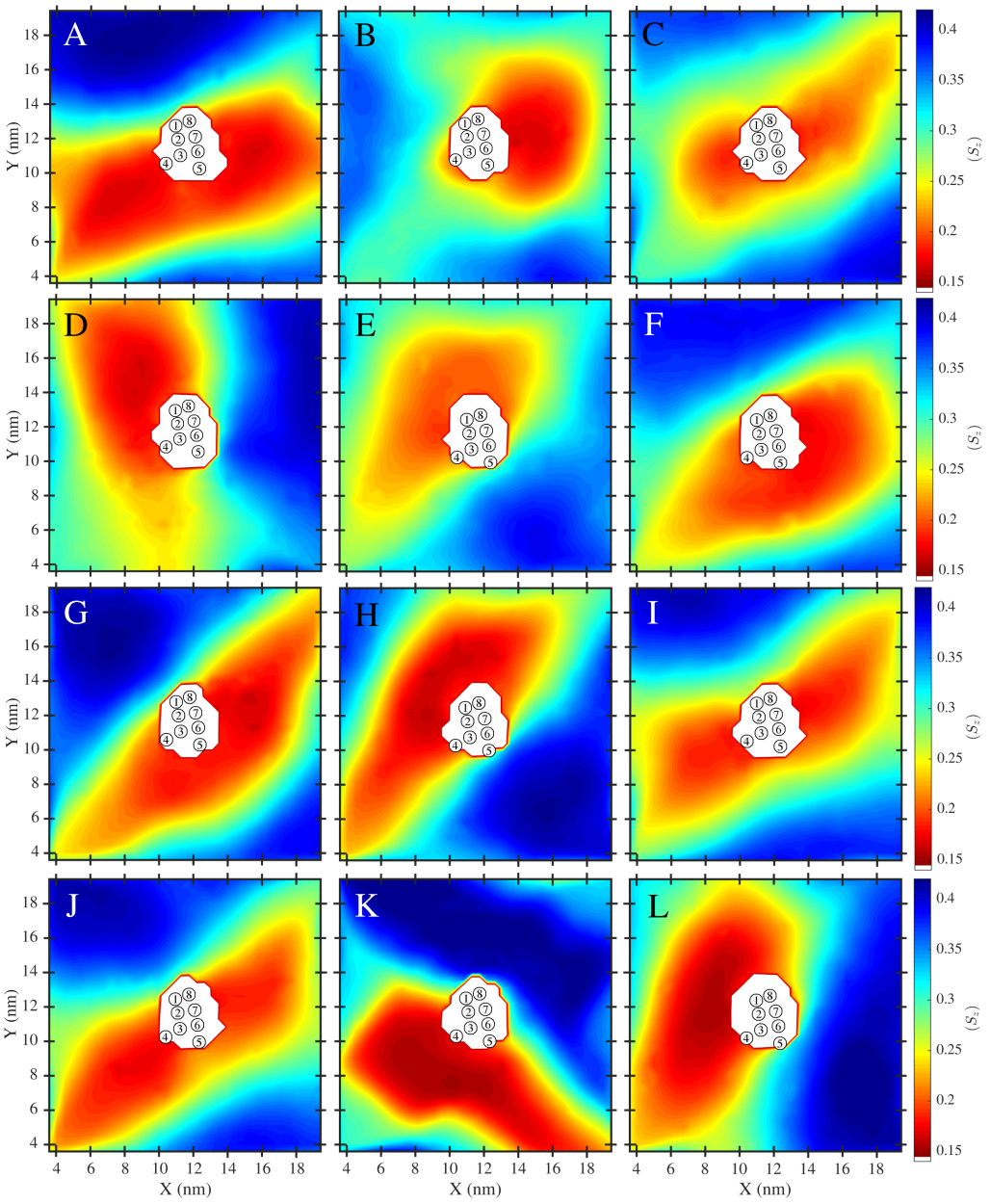


Figure F.4: The two-dimensional lipid chain order parameter in systems (A-J) R1-R10 and (K-L) RC1-RC2.

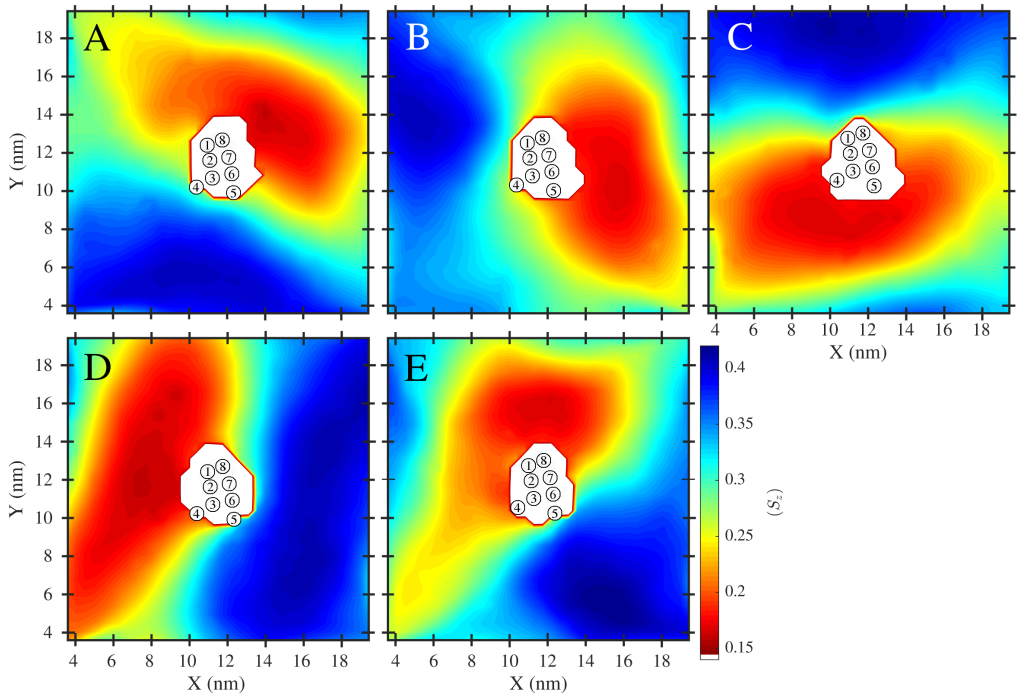


Figure F.5: The two-dimensional lipid chain order parameter in systems (A-E) O1-O5.

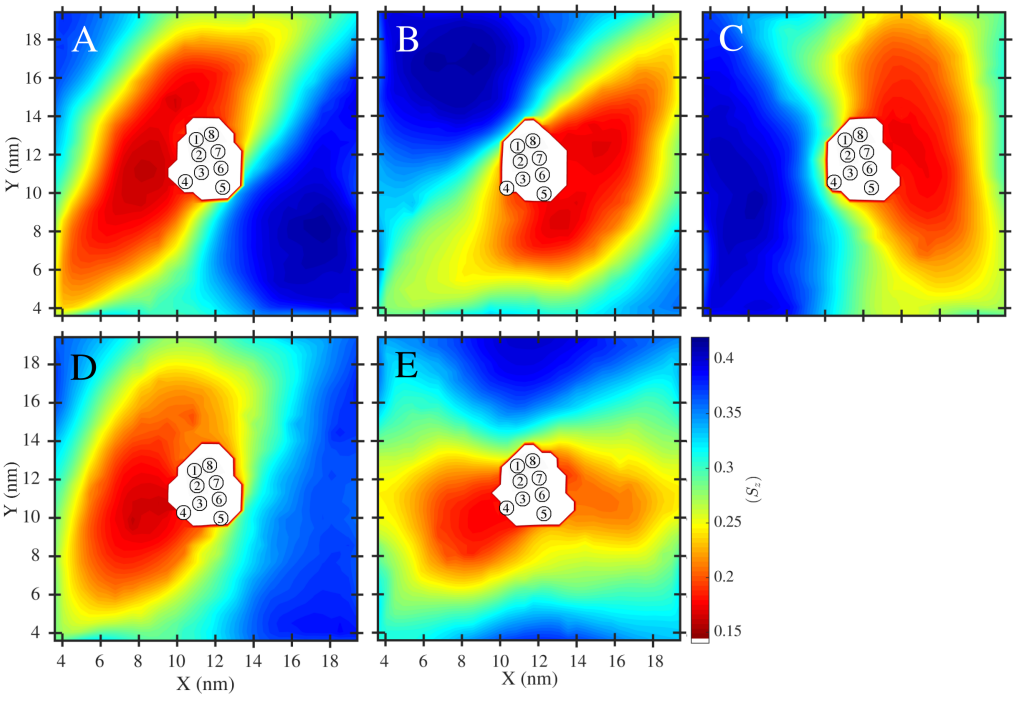


Figure F.6: The two-dimensional lipid chain order parameter in systems (A-E) D1-D5.